

**Culture and Expression in Mother-Infant Vocal Play: Do
Vowels Regulate Intersubjectivity?**

Niki Powers

PhD



The University of Edinburgh

2008

Abstract:

Three studies explored how vowel sounds are utilised by mothers and infants in the first year to regulate emotional expression. In Study 1, a cultural comparison was carried out. 6 English-speaking and 6 Japanese-speaking mother-infant dyads were filmed in their homes (3 male and 3 female infants in each country), when the infants were aged 4 months. Analysis was carried out of vowel sounds produced by mothers and infants and of bodily contact in 2 defined emotional situations ('engagement' and 'disengagement'). The findings presented here suggest that acoustic features of vowel sounds (pitch, intensity and duration) were found to be coordinated with bodily contact and correlated with specific emotional communicative contexts. Study 2 and 3 were only carried out in Scotland with English-speaking participants. In Study 2, an 'emotional voice' experiment was developed to test 11 infant's reactions to changes in pitch variation in mother vowel sounds, in specific emotional situations. In Study 3, 158 adult participants were asked to judge if isolated infant vowel-like calls (which had previously been coded for emotional content) expressed distinct emotions, and whether they felt any emotional response to the sounds. The findings suggest that acoustic features of vowel sounds appear to be utilised differently in 'engaged' and 'disengaged' interactions and in Japan and Scotland, suggesting that emotional communication develops in culturally specific ways. Results suggest that adults may be able to recognise emotional meanings in infants' early vowel sounds, and that infants often react in distinctive ways to emotional changes in their mothers' voices.

Contents

Contents.....	2
Index of tables and figures.....	5
Acknowledgements	7
Chapter 1: Introduction.....	8
1.1: Emotional Engagement in Early Vocal Communication, the Nature of the Problem.....	8
1.2: Setting the Scene: Theory for This Study.....	10
1.2.1 Principles of infant intersubjectivity and shared meaning.....	11
1.2.1.1 Imitation in infancy as intended communication.....	13
1.2.1.2 Proto-conversation and parental support for exchange of meanings	14
1.2.1.3 Active participation and emotional regulation.....	15
1.2.1.4 Cooperative Awareness - Sharing intentions by hearing feelings in time	17
1.2.1.5 Communicative Musicality: an integrative framework for description of human communication.....	19
1.2.1.6 Intersubjectivity and Cultural Learning	20
1.2.1.7 Making meaning in relationships: Intersubjective understanding of emotion and communication.....	22
1.2.2 Vocal Communication of Affective State.....	23
1.2.2.1: Human vocal communication	24
1.2.2.2: Expression, movement and experience in infant learning	27
1.2.2.3 Emotions and experience in relating with other persons	30
1.2.2.4 Emotional attachment, and development of feelings of the Self	33
1.2.3 Are Vowel Expressions linked to a General Intrinsic Motivation to Communicate?	34
1.2.3.1 Purposeful and imitative vocal action and feelings	35
1.2.3.2: The role of vowels in the early development of communication	36
1.3: The Research for This Thesis	39
1.3.1 The General Aims	39
1.3.2 The Specific Goals.....	39
1.4: Overview of the Thesis.....	40
Chapter 2: Study 1, Voices and emotions of mothers and infants in two cultures	41
2.1 Introduction: The Nature and Uses of Vowels.....	41
2.1.1. Vowels as a feature of communicative musicality that link emotion and Intersubjectivity	41
2.1.2. Prosodic phonology	43

2.1.3. The role of vowels in the vocal stream, and their function in communication with infants.....	45
2.1.4. Interpersonal 'contingency' in communication, and effects of its breakdown	47
2.1.5. The culture of emotion and vocal expression	49
2.1.6. Vowel features, communication of emotion, and learning the 'mother tongue' -- how Japanese and English cultures differ	51
2.1.7. Hypotheses of Study 1	54
2.2. Method.....	55
2.2.1. Participants and Procedure.....	55
2.2.2. Analysis.....	56
2.2.3. Acoustic Data.....	57
2.2.4. Analysis of Emotional Engagement.....	59
2.2.5. Analysis of Rhythmic Tactile Play	60
2.3. Results.....	61
2.3.1. Descriptive information	61
2.3.2. Acoustic features of vowel sounds	66
2.3.3. Emotional Engagement.....	69
2.3.4. Culture.....	70
2.3.6. Rhythmic/tactile play	71
2.4. Discussion	73
Chapter 3: Emotional Voices	80
3.1: Introduction: Mothers' Voice of Emotion and Infants' Listening	80
3.1.1: Connections between physical aspects of communication and emotional experiences.....	80
3.1.2: Timing, prediction, anticipation and narrative.....	82
3.1.3: Infants' perception of speech and characteristics of infant-directed speech	84
3.1.4: Vowels as a link between acoustic information and psychological and emotional experience	85
3.1.5: Infants' perception of vowel sounds – what features are salient and how are these features related to expression of affective state?	87
3.1.6: Vowel use in nursery rhymes and songs.....	90
3.1.7: Research Questions.....	93
3.2: Method.....	94
3.2.1: Participants.....	94
3.2.2: Procedure	94
3.2.3: Acoustic Analysis of the Mothers' Vocal Expressions	95
3.2.4: Rating of Non-Vocal Responses of the Infant to their Mother's Voice	96
3.3: Results.....	97
3.3.1: Dimensions of Emotion in the Mother's Voice	98
3.3.1.1: Pitch range of all vocalisations	98
3.3.1.3: Pitch Range of Vowels	104
3.3.1.4: Intensity of vowel sounds	105
3.3.1.5: Duration of mothers' vowels in each condition.....	106
3.3.1.6: Whispered vowels and infants' vocalisations.....	109

3.3.2: Infant reactions to mother's voice	110
3.3.3: Level of difficulty in 'acting' emotions.....	113
3.4. Discussion and Conclusions.....	116
3.4.1: Emotional characteristics of the mothers' vocal expression in the three emotional conditions.....	116
3.4.2: Rating of infant non-vocal reactions to the mother's voice.....	118
3.4.3: Levels of difficulty in 'acting' different emotions.....	120
Chapter 4: Can adults perceive emotions in infants' vowel sounds?.....	122
4.1: Introduction	122
4.2: Method.....	128
4.2.1: Emotions in Infant Vocalisations – Stimuli.....	128
4.2.2: Participants.....	130
4.2.3: Analysis	131
4.3: Results.....	132
4.3.1: Perceiving emotions in infant vowel sounds	132
4.3.2: Judgements about emotion.....	132
4.3.3: Effect of Sound files	134
4.3.4: 'Non-Matching' Judgements	136
4.3.5: How do adult participants describe infant emotions?.....	137
4.3.6: Adults' emotional responses to infant vocalisations.....	141
4.4: Discussion and Conclusions.....	144
4.4.1: Understanding emotions from infant vowel sounds	144
4.4.2: Adult emotional descriptions of and responses to infant emotional voices.....	147
Chapter 5: General Discussion and Conclusions	150
Appendix 1: Pitch plots of sound files	159
Appendix 2: Marwick and Trevarthen Category System for Affect	169
References	171

Index of tables and figures

Table 2.1: Participant Information	55
Picture 2.1: PRAAT Sound Editor Window	58
Graph 2.1: Pitch information for Japanese and Scottish Mothers and Infants.....	61
Graph 2.2: Intensity information for Japanese and Scottish mothers and infants	62
Graph 2.3 Vowel duration information for Japanese and Scottish mothers and infants	63
Graph 2.4: Frequency of vowel types for Japanese and Scottish mothers and infants	64
Graph 2.5: Frequency of vowels during ‘engaged’ and ‘disengaged’ communication	65
Graph 2.6: Vowel type in ‘engaged’ and ‘disengaged’ communication	66
Table 2.2: Pitch and Intensity for Vowel Types	67
Table 2.3: Descriptive information for acoustic features of vowel types	68
Table 2.4: Partial Correlations between Pitch and Intensity	69
Table 2.5: Acoustic features of vowels in relation to emotional 'engagement'.....	69
Table 2.6: Acoustic features of mothers’ vowels in cultural contexts	70
Table 2.7: Descriptive information for acoustic measures of mothers’ vowels during tactile play	72
Table 3.1: Mean length of rhyme in each emotional condition	96
Picture 3.1: One infant’s reaction to the ‘Sad’ condition.....	97
Table 3.2: Pitch range of all vocalisations, for each mother in each emotional condition (measured in Hertz)	98
Graph 3.1: Distribution for pitch range of all vocalisations in each condition.....	99
Graph 3.2: Pitch of one mother’s voice in the ‘Happy’ Condition.....	101
Graph 3.3: Pitch of one mother’s voice in the ‘Sad’ Condition.....	101
Graph 3.4: Pitch of one Mother’s voice in the ‘Bored’ Condition	101
3.3.1.2: Intensity range of all vocalisations	102
Table 3.3: Intensity range for all vocalisations for each mother in each emotional condition (measured in Decibels)	102
Graph 3.5: Distribution for intensity range of all vocalisations in each condition....	103
Table 3.4: Mean pitch range (in Hz) for vowel sounds for each mother in each condition	104
Table 3.5: Intensity measurements in Decibels (dB), for all vowels in each emotional condition	105
Table 3.6: Frequency of vowel type in each emotional condition	107
Graph 3.9: Vowel Types for each mother in the ‘Happy’ Condition	108
Graph 3.10: Vowel Types for each mother in the ‘Sad’ Condition	108
Graph 3.11: Vowel Types for each mother in the ‘Bored’ Condition	109
Table 3.7: Whispered vowels in each condition	110
Table 3.8: Concurrent infant vocalisations in each emotional condition.....	110
Table 3.9: ‘Matches’ for each condition	111
Table 3.10: Total number of codings given by each rater for each emotion (including both ‘Matches’ and ‘Non-matches’)	112
Table 3.11: Total ‘Matches’ (out of 9) for each infant	113
Table 3.12: Mothers’ rating of level of difficulty in ‘acting’ emotions.....	114
Table 3.13: Frequency information for emotional condition and mothers’ rating of difficulty.....	115

Table 4.1: Infant Age, Sex, Description of Vocalisations and Emotional Message..	129
Table 4.2: Participant sex and parental status	130
Table 4.3: Descriptive information for sex, parental status and judgement type	133
Table 4.4: Summary of MLS analysis of participant group as a predictor of judgement type.....	134
Table 4.5: Frequency Information for Sound Files by Judgment Type and Participant Group	135
Table 4.7: Summary of MLS analysis of sound file, as predictors of ‘Matching’ judgement type.....	136
Table 4.8 ‘Non-Matching’ Judgements	136
Graph 4.1: Description Type for Females and Males	138
Table 4.6: MLS Summary for Description Type	139
Graph 4.2: Judgement Type by Description Type	140
Table 4.7: MLS Summary for Judgement Type and Description Type.....	140
Graph 4.3: Response Type for Females and Males	142
Graph 4.4: Response Type for Parents and Non-parents	143
Graph 4.5: Response Type by Judgement Type	144
File 1: Ho1, 23 week old female infant, positive emotion (‘uh’ sound).....	159
File 2: Ho2, 23 week old, female infant, negative emotion (A shrieked ‘a’ sound – scratchy sound so pitch couldn’t be measured completely).....	160
File 3: Is1, 23 week old, male infant, positive emotion (A shrieked ‘ee’ sound).....	161
File 4: Is2, 23 week old, male infant, negative emotion (An ‘eu’ sound)	162
File 5: Jan 1, 28 week old, female infant, positive emotion (An long drawn out ‘owue’ sound).....	163
File 6: Jan 2, 28 week old, female infant, negative emotion (An ‘a-eu’ sound).....	164
File 7: Ai 3, 36 week old, male infant, positive emotion (An ‘a’ sound)	165
File 8: Ai 1, 36 week old, male infant, negative emotion (‘ue’ sound)	166
File 9: Li 1, 46 week old, female infant, positive emotion (A ‘ya’ sound).....	167
File 10: Li 2, 46 week old, female infant, negative emotion (An ‘oh’ sound)	168

Acknowledgements

This has been a long journey, sometimes through difficult waters, sometimes through beautiful scenery. I want to say thank you to the people who have supported, encouraged and inspired me on the way. To my husband Rhys and my sons Stevie and Joe, I love you. Thank you for your patience, your understanding, your constant support and your practical help. To Colwyn Trevarthen and Morag Donaldson, thank you for going beyond the call of duty. I feel lucky to have had your support. You have both given me a gift that I will carry with me for the rest of my life. I would like to thank my Japanese colleagues and friends Koichi Negayama, Miwa Hirose and Noriko Kawahara and all the mothers and babies who participated in my research. Thank you to George Tait and Bob Morris, I miss you. Thank you to David Wilkinson, Jimmy Duncan, Bill Robertson, Roy Welensky, Anna Sim and Katie Keltie for all your kind help and support.

Chapter 1: Introduction

1.1: Emotional Engagement in Early Vocal Communication, the Nature of the Problem

My thesis research is planned to test the idea that vowel sounds produced by mothers and infants when they are in communication informs them with some sense of their own and each others' emotional states. I ask an apparently simple question: "Are vowels a medium through which emotional communication is expressed and understood in the first year of life?". To put this question in a natural context, several aspects of communication and motivation need to be considered. A review of the literature will present strong evidence that infants actively engage with communicative partners to regulate affective state and maintain mutual involvement in an engagement of motives, and it appears that all modalities of expression and perception may contribute. The above research question must, therefore, be grounded within a theory of the core integrative psychological processes that would be necessary for such a multi-modal regulation to occur within and between any two human subjects. It will assume that the following five conditions or factors in regulation of motive states must be considered in any attempt to explain how intersubjective communication can develop in infancy:

1. Changes in the *physiological state* of arousal and alertness of the infant.
2. The *seeking motives* in the infant and caregiver, coupling impulses in each of them that organise their movements and facilitate selective awareness by focusing and directing attention.
3. The *stimulation of interest* in the infant necessary for learning new ways to communicate.
4. *Attachment emotions*, in both infant and adult that express affectionate concern for well-being, or states of pleasure or distress, in the other.

5. An additional *Companionship Motivation* promoting shared interest and pleasure in activities oriented to the experience of objects and events in the present world, permitting the performance of joint intended actions, or tasks.

The complexity of human communication is challenging. There are advantages in a research strategy that isolates one feature of behaviour or element of communication for detailed examination. At present the underlying processes that mediate emotional communication by non-verbal means are not fully understood and it is uncertain if such a single feature of communication, such as a vowel sound, could provide a signal for two humans, one of whom is very young, enabling them to share and respond to emotional and motive states. However, the special salience of long vowel-like vocalisations in mother-infant communication everywhere they have been studied encourages the belief that these sounds are a natural, and possibly essential, component of their mutual awareness.

This review will consider how researchers have understood acoustic aspects of mother-infant interactions, in particular how they have described the vowel-like sounds made by the infant and the parent. The literature provides detailed descriptions of the many aspects of behaviour that play a conspicuous role in communication and emotional engagement between parents and infants. In this research project the non-vocal communicative processes will be considered in less detail, but it is assumed that vowel use is only a tiny piece of the puzzle of human sympathetic communication. Further detailed examination of other aspects of communication, both phonological and gestural, will be required before we can more fully understand how acoustic features of the voice contribute to our psychological experiences of emotional expression and perception.

First, evidence will be summarised relating to a range of cognitive, emotional/motivational and motor theories of how human communication begins and how it develops in the period before language. Accounts of the earliest communicative abilities are reviewed, to consider if a theory of Innate Intersubjectivity may define parameters that can be used to trace the steps of normal

communicative development. A challenge in the current research is to find a perspective by which the relationship between measurable physical (acoustic) properties of expressions and the inferred causal psychological processes can be clearly understood. The theory of Intersubjectivity claims to explain the process of human communication in terms of psychological sympathies involving all the fundamental control processes of conscious agency and investigators seek evidence that infants and mothers can indeed engage with each other consciously, intentionally and emotionally from birth.

The review will, then, weigh the evidence for the idea that vowel sounds may play a central role in the complex inter-mental processes that are presumed to take place during communicative interactions between mothers and infants. An understanding of how acoustic features of extended vocalisations are utilised by mothers and infants during development will gain us insight into how this one element of expression is integrated in the functioning of multimodal perceptual and motor systems that are united in the rhythms of communicative behaviour. It will look further at how the communication of emotion can be changed by experience, exploring how motivation and culture influence development of affective vocal expressions. To test how communicative processes may be influenced by cultural factors, a comparison will be made of developments in communication between mothers and infants in Scotland and Japan.

1.2: Setting the Scene: Theory for This Study

To provide a framework for my research, this section will define the terms used in studies of communication with infants and will examine theories of how intersubjectivity can be regulated. The research questions that will be set out in 3.1 have been framed in the light of the knowledge presented in this review.

1.2.1 Principles of infant intersubjectivity and shared meaning

From birth, infants manifest a desire to communicate with other humans. It is useful at this point to state a working definition of what is meant here by the term ‘communication’. In psychology it is generally assumed that communication may or may not involve consciousness. Buck (1984), for example, defines communication on two distinct levels. He uses the term ‘spontaneous’ to define expressive behaviour that allows animals to communicate by “externally visible and accessible manifestations of internal states...based on a biologically shared signal system” (p. 6). Interactions between even highly evolved animals are generally considered to be at this level.

Buck defines the second ‘symbolic’ level of communication, unique to human beings, as ‘intentional’ where “communicative behaviour has an arbitrary, socially defined relationship with its referent...knowledge of which is shared by sender and receiver” (p. 7). Buck relates symbolic communication specifically to language, and this presents some problems in for interpretation of infant pre-linguistic communication, which has features of both of these levels of description. However, he admits that he has not fully explained the nature of the relationship between spontaneous and symbolic levels of communication, and it is important that Buck accepts that culture is the product of conventions that are built through interactions mediated by spontaneous actions in communication. To understand the developmental origins of symbolic communication and language we must consider how conventions are learned by sharing of intentions and interests in early childhood, when words are not yet understood by the child.

Research over the last few decades has found that infants respond in a coordinated and directed way to communicative expressions of persons who engage with them (Bateson, 1971; Stern, 1985; Locke, 1993; Beebe and Lachmann, 2002). Infants also demonstrate that they are aware that they are the focus for the communicative behaviour directed to them. They are sensitive to shifting emotional states that underlie communicative behaviour from others and they have the ability to express their own motivational and emotional feelings. This provides evidence for the theory

that infants possess innate Intersubjectivity which leads on to cultural learning (Trevarthen and Hubley, 1978; Trevarthen, Murray and Hubley, 1981).

Trevarthen et al (1981) describe two distinct phases in the development of communication with infants, characterised by new interests and developing motor abilities. Primary Intersubjectivity is evident during the first few months of life and is driven by the infant's motivation to attend to and interact with their closest caregivers. At this stage, the infant can be described as possessing "an active and immediately responsive conscious appreciation of the adult's communicative intentions" (Trevarthen and Aitken, 2001, p. 5). Secondary Intersubjectivity is characterised by the infant's growing interest in the wider environment and the things contained within it. Around the 9th month of life the infant begins to perceive their communicative partners as "a source of new ideas concerning objects" (Trevarthen, 1994, p. 235). At this point the infant will both display more praxic or 'object using' behaviours and pay attention to the ways other people use objects; their communicative development expands to include interest in objects as a shared point of reference in 'joint attention' (Trevarthen and Hubley, 1978; Tomasello, 1988; Legerstee, 2001).

Trevarthen defines the overall process behind intersubjectivity in infancy as follows:

"Infants try to engage with the consciousness of other persons, and they offer significant messages in the ways they regulate orientation to events around them in response to other persons' interests. They make efforts to share experiences and purposes and, before many months, they take up other persons' meanings through communication with their feelings and interests, fitting into the activities of others' minds by an imitative sensitivity to the expressions of their bodies."

(Trevarthen, 1994 p. 219).

Several strands of evidence that support this interpretation of infant communication will now be explored.

1.2.1.1 Imitation in infancy as intended communication

There has been intense debate about the importance, and indeed the existence, of neonatal imitation. Moreover, there is no one 'standard' definition of the term 'imitation'. Within a theory of Intersubjectivity, imitation could be defined as, 'a communicative behaviour, where one communicative partner copies or matches any behaviour (verbal or non-verbal) directly or indirectly in response to a similar behaviour by their interactive partner, while they respond to one another as human agents'. It is assumed that the imitation is motivated *as communication*.

The purpose and nature of early imitative behaviour is disputed, because it has not been made clear what the motivations for this type of behaviour are. Some researchers claim that imitation allows humans to develop a subjective self through emotional, embodied attention to others' actions, thus creating or facilitating intersubjective intimacy (Nadel, 2004; Zeedyk, 2004; Nagy, 2004). However it develops, and whatever its role may be in acquiring a Self, it is now beyond dispute that imitation takes place from birth (Kugiumutzakis, 1993; Heimann, 2001; Nagy, 2004). Furthermore, there is recent evidence from neurophysiological studies that 'mirror neurons' in the brain provide the biological basis for imitative intentions in humans and other primates (Rizzolatti and Arbib, 1998; Rizzolatti and Craighero, 2004).

Donald (2001), who identifies 'mimesis' as the primary form of message that led to the evolution of human symbolic narration, describes the early development of imitation as 'interlinking the infant's attentional system with those of other people' (p. 255). He suggests that through mimesis of communicative actions, infants can formulate participative routines that become more complex over time. Donald asserts that for an infant to learn human communication, this reciprocal imitation is an essential foundation.

Recent research carried out by Nagy proves that neonatal imitation is not a 'reflex' behaviour' or automatic reaction to stimulation from the other person (Nagy and

Molnar, 1994; 2004; Nagy, 2004). It was found that infants both imitated and 'initiated' or 'provoked' communicative gestures. Their imitative actions were purposeful and observant. Measurement of infants' heart rates whilst they were involved in reciprocal 'imitation cycles' found that when the infant was *imitating*, with effort, their heart rate increased, and when they were *initiating* or *provocating* their heart rate decreased, as the infant was waiting attentively for a response. The suggestion that infants were literally *waiting* for a response was backed up by the fact that initiation was only coded if the infant was also looking directly at their communicative partner, watching them. Nagy (2004) proposes that these cycles of imitation and initiation are "the first turn-taking dialogues" (p. 11). The evidence does seem to prove that communicative abilities required for intersubjective exchanges are in place at birth, and to offer an explanation how they could be built upon during the ontogenesis of communication.

1.2.1.2 Proto-conversation and parental support for exchange of meanings

As the baby's innate communicative abilities are fostered and supported by affectionate caregivers, and the infant's expressive and perceptual faculties undergo rapid development, other features that are central to the maturation of intersubjective interaction begin to flourish. Early 'proto-conversations' between infants and those around them begin to take shape. The anthropologist Mary Catherine Bateson first defined 'proto-conversation' in the 1970s from observations she made of interactions between a 9-week-old infant and the mother, and since her original proposition that this behaviour is a foundation for learning of language and 'ritual practices', researchers have confirmed that these early conversations begin to be well-regulated at around 2 months (Trevvarthen, 1979; Horner and Chetnik, 1986; Rochat, 2001a; Gratier, 2003; Reddy and Trevvarthen, 2004). As the infant develops an increasingly critical consciousness, the repeating 'rituals' of infant-parent proto-conversations become more complex and more tightly organised in timing and structure (Rochat et al, 1999).

These early conversations between infants and their parents are structured in a similar way to conversations between adults (Stevenson et al, 1986; Jaffe et al, 2001). They

contain features of 'turn-taking', 'mutual response and initiation', and 'co-ordination of gaze' with facial and bodily movements and hand gestures (Jaffe et al, 2001; Reddy and Trevarthen, 2004). Gratier (2003) points out that these early conversations also take place within specific cultural contexts and, even at this early stage in development, infants are integrating a preferred cultural rhythm and style into their communication. At the moment, it is not entirely clear exactly how this complex coordination of inherent and acquired 'habits' of communication takes place. Language simulation studies have suggested that young infants can discriminate between languages based on rhythm classes (Ramus et al, 1999). It is likely that coordination of rhythmic features of early proto-conversation helps the infant to form expectancies about paralinguistic features of verbal communication, including the *timbre* or quality of voice that gives utterances emotional appeal.

Of course conversations require the cooperation of two persons. The infant's main caregivers provide essential 'symbiotic' support to the way that these early interactions provide emotional satisfaction for the infant, and how they will form a secure basis for the infant's future cognitive development (Valsiner, 1988; Papoušek and Bornstein, 1996). When mothers and fathers engage with infants, they do so in a systematic, predictable and responsive way. They assign meaning to the vocalisations produced by infants and they also try to interpret the intentions of their non-vocal behaviours (Trevarthen, 1999). It may also be that the simple, repetitive structure that is so typical of proto-conversation allows the infant to tie emotional meaning to the motivated elements of parents' talk with them. When caregivers take part in protoconversation with infants, they have to engage on an emotional level with the infant for the interpersonal process to be successful (Reddy and Trevarthen, 2004).

1.2.1.3 Active participation and emotional regulation

Infants appear, within a few months of birth, to be skilled at participating in high affect communication in games and songs (Stern, 1974b; Bruner and Sherwood, 1975; Stern, 2000; Trevarthen, 2001). As well as being a necessary prerequisite for the healthy development of intersubjectivity, active participation in interaction provides evidence that infants are dynamically involved in the process of communication and its 'ritualisation'. On the other hand, infant withdrawal from communication is

commonly associated with a range of clinical problems. Guedeney and Fermanian (2001), who have developed an assessment scale to measure infant withdrawal, state that brief periods of withdrawal are normal for all infants, but longer periods of emotional and communicative withdrawal reliably indicates physical and/or emotional problems.

To be able to participate and withdraw from communicative involvement in a coordinated and appropriate way an infant must be more than a reflex system. Infants have the ability to actively *regulate* their affective interaction with other persons, and active participation in communication is an indicator of healthy development. The infant's participation in communication is encouraged when their partner is *engaged* with them fully, is *attuned* to temporal aspects of the infant's vocal and non-vocal communicative behaviour, and responds *contingently*, with immediate temporal coordination (Papoušek, 1992; Papoušek and Bornstein, 1992; Jaffe et al, 2001; Reddy and Trevarthen, 2004). If an infant's communicative attempts are not supported (for example, by a mother who is suffering from postnatal depression), the infant will still strive to make communicative contact (Gratier, 2003). This confirms an active participation by the infant as well as the emotional need for active participation of both communicative partners.

Nevertheless, some psychologists doubt that prelinguistic infants, and especially before the age of 6 months, have the capacity to actively participate in interactions with others. Some argue that infants acquire the ability to actively participate because they are treated, by parents and other caregivers, 'as if' they are intentional beings (Dunn, 1999; Reznick, 1999). It does seem to be the case that caregivers treat infants as if they are acting in an intentional manner (Bruner, 1983; Papoušek, 1992), but this does not necessarily mean that this attitude or the belief it expresses is the *cause* of the child's developing ability for intentional action. It is difficult to understand how infants could coordinate a whole range of complex gestural and vocal aspects of expression in an appropriately responsive way with another person (Jaffe et al., 2001) if they did not have some means by which to actively direct their involvement, albeit in an immature and developing way.

1.2.1.4 Cooperative Awareness - Sharing intentions by hearing feelings in time

If the infant's active participation in communication is accepted, then we can proceed to find an explanation for another important development in early intersubjectivity, which is the appearance of cooperative awareness and sharing of intentional states, and of emotions associated with intentions. Sharing intentions is an adaptive function of human intersubjectivity, essential for learning meaning in an artificial culture, and it explains why coordination of actions and feelings is so precociously developed in infants. Coordination of information about the objective world with signals that express emotions requires a further level of motivation and awareness of cooperation to allow mutual understanding of what information 'means'. The development of communication is evidently based upon the sharing of motivational and affective states (Stern, 2000; Papoušek and Bornstein, 1992; Panksepp, 2001), and this dynamic intersubjective process leads to the development of cooperative awareness. Mothers and infants interacting demonstrate a second to second awareness of the motivational states of their communicating partner (Papoušek, 1992; Wittmann and Poppel, 2000). The important question here is by what means do infants perceive and express motivational states from the time they are born. It is likely that key aspects of motivation are expressed and perceived by the utilisation of temporal and acoustic features of the voice.

Jaffe et al (2001) carried out a comprehensive study into the co-ordination or mutual regulation of affect in infants' communication with mothers and female strangers in different environments. This research examined vocal rhythms in communication using time-series regression and found that vocal interactions are coordinated on an interpersonal level and that measures of coordinated interpersonal timing (CIT) are related to later security of attachment between the infant and adult. Trehub and Nakata (2002) reviewed evidence showing that vocal communication is an important channel by which emotional states are shared. Crown et al (2002) found that 6-week-old infants coordinated their gaze with the vocalisations of their mothers, and Horner and Chetnik (1986) found similar results with 18 week-old infants. This indicates that even in the first few months, infants are attentive to acoustic features of the voice and

are encouraged by hearing vocalisations to maintain their awareness, to track changes and to synchronise their own communicative efforts in different channels with those of their partners. It is confirmed that the process is naturally multimodal, involving the coordination and awareness of a complex array of affective expressions.

However, awareness of affective states is only part of what is required for communication. Schögler (2003) proposes that the sharing of “motivated psychological time is the key to social communication in all animal species” (p. 9). To be aware of gestures and to coordinate one's own gestures with the gestures of others does indeed require coordination of timing. There is a great deal of published evidence showing that an ability to coordinate temporal aspects of communicative expressions, both vocal and non-vocal, with expressions of a partner is evident from birth. Control of timing and rhythm is essential for an infant to be able to systematically monitor and organise their affective experiences when communicating with others (Byers, 1975; Fernald and Mazzie, 1991; Morgan, 1994; Beebe and Lachmann, 2002). By coordinating with the temporal patterning of another person's behaviours, one may create a shared time of moving, so that each can give and receive continuous feedback to regulate emotional aspects of the interaction.

Some researchers suggest that *matching* innate rhythms may allow the infant to share representations in the brain that facilitate the development of music and language systems, thus setting the foundations for emotional, social and cognitive development (Ramus et al, 1999; Ramus et al, 2000; Panksepp and Bernatzky, 2002; Hauser and McDermott, 2003). Others focus on the importance of *contingency* of responses in rhythmic exchanges of cooperative awareness and the sharing of emotional states (Stevenson et al, 1986; Rochat et al, 1998; Nadel et al, 1999; Gratier, 1999). As infants interact with their caregivers, mutual awareness and sharing of intentions seem to rely on emotional expectancies being met in experienced time. It may be that particular physical properties of the information that is exchanged in communication help to form representations in the brain that allow these emotional expectancies to be developed.

This idea links to recent theories to explain innate *intersubjective sympathy*, which propose that the sharing of a mutual time of action and experience is made possible through matching of temporal aspects of voice and other motor activity generated by human brains, a matching that is prominent in interactions between infants and their caregivers. This mutual awareness and sharing of the rhythms of affective states would lead infants, as they share vocal and non-vocal communicative gestures with more experienced others, to develop an understanding of the complex interpersonal and cultural meanings that are created in cooperative human communities.

1.2.1.5 Communicative Musicality: an integrative framework for description of human communication

Malloch (1999) defines 'Communicative Musicality' in terms of 3 components - 'pulse', 'quality' and 'narrative'. His acoustic analysis demonstrated that a 6-week-old can enter a rhythmic communicative partnership with a parent, with characteristic syllable and phrase elements, and he showed important developments in timing and expression in the first 6 months. A vocal interaction of a two-month premature baby with the father demonstrated the same fundamental features. Parents and infants mutually adjust the elements of 'pulse' and 'quality' to produce a 'narrative', in 'conversation', anticipating and regulating the cycles of emotional intensity, which Stern calls 'proto-narrative envelopes' (1985). Jaffe et al (1999) measured 'time vocalising', 'pause' and 'switching pause' parameters to show how adults 'infantise' speech timing when talking to an infant. Vocal timing of both adults and infants changes when they converse with 'coordinated interpersonal timing'. Thus infants can actively regulate their behaviour in time to respond to parents' vocal emissions. Quality of vocalisations was not analysed. Infants also attend to tonality, harmony, frequency, pitch and melodic contour of maternal vocalisations, and it is likely that many of these features may be discriminated in the last 2 months of gestation, in utero (Locke, 1993).

'Musicality' has been found to play an important part in the development of communication (Papoušek, 1996). Even before birth, infants are exposed to humanly organised acoustic information that contributes to the development of auditory architecture and allows the infant to be born with preferences for and memory of

particular features of speech (Lecanuet, 1996; Abrams and Gerhardt, 1997). The musical nature of infant-directed speech provides a perceptually attractive and functionally useful means by which infants can begin to segment the speech stream. Musicality in vocal behaviour of infants' companions provides 'templates' that may aid the infant to predict, anticipate and remember repeating narrative patterns. There is research showing that infants are sensitive to changes in many aspects of voice, such as pitch, intonation, timing and rhythm (Fassbender, 1996; Trehub, 1990; Trehub, Schellenberg and Hill, 1997).

1.2.1.6 Intersubjectivity and Cultural Learning

Humans share practical and emotional meanings through acceptance of the conventions of culture. As Bruner (1996) puts it, culture 'shapes the mind' because human minds do not develop alone. They grow and form within cultural landscapes, first in an intimate and private culture that is interpersonal in the family, then in a wider, praxic and social setting. The emergence of cultural learning appears to begin with early intersubjective communication, where infants are trying to make sense of the world around them by joining in cooperative play. Bruner proposes that culture gives meaning to the infant's experiences and interactions within the community of which she or he is a part.

Merlin Donald (2001) advances this argument by suggesting that humans are conscious of their experiences because they are intersubjective, because they can engage with other people. He states:

“A key step in this process is the interlinking of the infant's attentional system with those of other people. Reciprocal eye contact, or gaze, is one of the channels through which this develops. Voice and touch are also common channels. These are all involved in familiar circular routines, such as greeting, hugging and playing. Moreover they must become reciprocal, so that the infant is an active participant in these routines. These interactions are learned by the infant primarily through intimate personal encounters, starting with early mutual imitations with the mother and broadening to include more

elaborate exchanges of facial expression, voice and gesture, which lead to play acting and various social games. Such encounters serve to build a complex repertoire that will regulate shared attention for the rest of the child's life. The cultural repertoire constitutes a scaffolded system, with each new level adding to a vast cultural edifice of control. The result is to interlock the infant's growing mind with those of its caretakers and ultimately the broader society.” (p. 255).

Different cultures may have very different beliefs about childrearing and about appropriate emotional expression, and these may affect infant development from early stages. For example, Suzuki (2000) found that when Japanese adults were asked which values they consider to be important in childrearing, they emphasised 'empathy', forming relationships with others and working within a *group*, whereas in western cultures the importance of the *individual* was highlighted. These ideas conspicuously affect how feelings between persons of any age are expressed in the two cultures. Such differences in 'theories and practices of child development', beliefs that are pervasive throughout each society, condition how parents bring up their children, how they expect emotions to be expressed and how children go on to be educated (Bower, 1998; Weisner, 2001; Diener and Lucas, 2004).

Differences of this kind are also evident in much smaller communities or subcultures, and, given the influences of individual infant personalities and differences between individual parenting styles, intimate private cultures will be more diverse still. However, in all cultures and in all families emotions are expressed and communication occurs. In the learning of the culture of interpersonal manners, as in the learning of language, innate physical and psychological communicative systems provide the motivation and the regulation for the creation of an infinity of meanings by a relatively small number of vocal and non-vocal gestures and expressions.

1.2.1.7 Making meaning in relationships: Intersubjective understanding of emotion and communication

There has been much debate as to how emotion and communication of meaning interact. According to Dillard (1998) emotional processes and cognition are inter-related and each can influence the other. Sally Planalp (1998) asks what is ‘understood’ when emotions are expressed. Communication affects how we perceive and express feelings; it is both the cause and product of emotional experience and of motivation to make sense of the world as other do, and to assert our place in it. Communication is action that is shared, and it transmits motives between communicative partners. But it is still unclear in what way particular measurable physical features of communication behaviour are linked to emotional motivation and the mechanisms of expression.

Feelings and motivations are expressed and perceived through multimodal channels of movement and perception. The psychotherapy literature confirms the therapeutic value of establishing the bodily location of feelings, precisely because ‘mental’ emotions are seen as intrinsically linked to ‘physical’ experience in the body (Loughlin, 1999). Anderson and Guerrero (1998) suggest that, while there is little doubt that social life creates emotions that are generated through interpersonal interaction; we are still unclear about how physical features of interaction are translated into emotional meaning. Papoušek (1996) states that vocal communication provides affective information at three levels: the expression of a specific emotion in the individual, the identification of contextual information, and “an appeal addressed to the social environment” (p. 41). In this theory, the emotion provides the motivation to communicate with another person, with the aim of eliciting a response and perhaps a confirmation of existence. In other words an emotion needs to be expressed, to be about something, and to address someone. Papoušek concludes that, as an adaptive feature of the comparatively complex human brain, communication of emotions in these three ways has evolved to contribute to the forming of human cultures.

Donald (2001) formulates a theory of the communicative process in bio-cultural terms. He describes human communication as taking place within a collective consciousness

in the following way: “The machinery of language is far too exquisite to have been encoded entirely inside something as unpredictable as the genome of a developing brain. It was more probably shaped by the demands of a communicative universe that was much larger than one contained inside a single brain and was instead provided by a community of brains...a cognitive community” (p. 253). In these terms, communication is a process of sharing experience by linking with other communicators, and this requires communicators to express, perceive and regulate emotional states.

For the purpose of my research, developments in infancy communication and emotion will be assumed to be functionally interdependent. Communication will be defined as expressive and perceptive actions driven by motivations and emotional regulations that are both spontaneous and potentially symbolic in nature, operating and growing within a socio-cultural framework.

Theories of motives and intentions and their relation to awareness, emotions, and the acquisition of cognitive functions -- how dynamic, action-related consciousness is shared -- are required to understand the development of symbolic communication. Donald (2001) rejects the theory that humans are specifically ‘hardwired’ for language, proposing rather that a child can acquire this, and other, symbolic means of communication through their propensity for awareness and sympathetic response to rhythm and mimetic memory expressed in the behaviour of social companions. He suggests that sharing communication is based on “kinematic imagination...the ability to envision our bodies in motion” (p. 271). This process, the source of mimesis, allows infants to represent actions, to remember patterns of activity, and to carry out and refine their own actions. The emotional aspect of this process is the driving force and the factor that gives meaning to the actions that humans communicate about.

1.2.2 Vocal Communication of Affective State

Communication is, at all stages of life, the transmission of information about both ‘purposes and concerns’ (Donaldson, 1992). It exercises both cognitive and affective activities of the mind. Central to the process is the expression and perception of

emotions, showing what you feel, perceiving the feelings of others and interacting within a context driven and shaped by giving and receiving feelings. Visible, non-vocal features of communication by facial expression and bodily gesture are much studied and have been proved to provide affective meaning in communication (Ekman, 1972). Hobson (2002) describes how humans have “a universal body language, more basic than the language of words that connects us with other people mentally” (p. 48). However, the literature is less clear about how acoustic aspects of vocal interaction, other than the words of speech, provide a medium through which mental meanings can be shared. Recently research on infancy has led the way to a better understanding of emotional expression in the voice.

Research shows that many acoustic properties of infant directed speech convey emotion (Kitamura and Burnham, 1998; Trehub and Nakata, 2002). The development of emotions in infants is usually observed to gain information about self/other regulation and attachment. To understand how acoustic features of vocalisations and in particular how vowels are utilised in all kinds of intersubjective communication, it is necessary to look at the wider picture, including the development of shared interest of circumstances in the habitual world. Once a picture of the purposes of communication is established, it is easier to focus back down to the micro level to gain some understanding of how variation in vowel use might contribute. Then it will be possible to make some theoretical judgement about the role that vowels play in intersubjective communication of emotions.

1.2.2.1: Human vocal communication

Human emotional communication is uniquely complex, but there are many similarities between the ways animals and humans convey emotional messages that are important in the regulation of their social life. An examination of the nature of animal communication can give insight into possible evolutionary explanations for the development of more complex human communication. This can provide a background for a theory of the functions of communicative signals and how human signals might have evolved.

Donald (1999) suggests that the main advance that enabled humans to develop their communicative abilities beyond those of apes and chimpanzees is the acquisition by human brains of a new capacity to fine-tune motor skills. Donald points out that humans (even very young children) practise their movements solely for the purpose of improving them – animals do not do this, except in play which is limited in its purposes, being important for exercise of the body and alertness and for establishing social positions for individuals. Donald suggests that experimentation with action allows humans to evolve much more complex cognitive abilities. Importantly for the present research, he states that this development is built upon a uniquely complex temporal patterning of acoustic expressions and gestures that are shared in social interactions from infancy.

Human infants appear to be born with a level of communicative ability that has not evolved in other social animals. Chomsky (1980), as a theoretical linguist, observed that human children have an apparently limitless ability to acquire language, and he inferred that humans have an innate ‘Language Acquisition Device’ (LAD), one that equips children with language universal cognitive processes that are activated when other people communicate with them. Chomsky (1996) later refined his theory to describe this device in the broader terms of a ‘language faculty’, suggesting that this is a biological feature, which can be shaped by experience to allow the internal representation of phonetic and semantic information, eventually leading to formal language. He and his colleagues later argued that the faculty of language could be classified into a ‘broad’ faculty, which functions within a wide range of sensory-motor and conceptional-intentional systems, and a ‘narrow’ faculty that is specific to humans and which is responsible for ‘recursion’ in language (which they refer to as a “domain of communication” (Chomsky, 1996, p. 1569). Hauser et al (2002) state that this is what allows human infants to create an infinite number of possibilities from a limited number of fundamentals. This theory suggests that infants must learn language automatically.

Donald (2001) differs, arguing that, although an innate ability that facilitates language acquisition, the emotional capacity to actively engage with ‘cognitive communities’,

mentioned previously, lays the foundation for the seemingly limitless complexity of communication in all its forms. It is this process that allows infants to have an awareness of messages in talking a long time before they can speak.

Birds and other animals produce and react to emotionally specific vocalisations, but the process of learning how to vocalise so other individuals will respond appropriately is limited. The vocal signals of birds carry information about the motivational states of the signaller (Zahavi, 2002) and there are associations between specific patterns of bird song and the many communicative functions. Zahavi believes that vocal signals provide information about posture and the level of tension, and that this corresponds in a functional sense to messages provided in the non-verbal communication of humans.

Vocal signals of animals convey information about how the emitter is moving, or how they intend to move. In particular vocal signals provide information about the vocaliser's level of motivation and assessment of risk. Regular patterns of sound are recognisable as warnings or threats, and some animals and birds will imitate the sounds they hear others make. But there is more to symbolic semiotic communication of humans than complex repertoires of acoustic behaviour and imitation.

Manipulation of acoustic information does not become symbolic by quantitative increase of the signalling powers that animals show. It must have a *negotiated meaning* -- emotional and motivational significance that is shared and created in a social context of a more elaborate kind.

Symbolic meaning must be processed by a brain capable of understanding that this meaning is shared in an arbitrary way, and has significance because it has been so shared. According to Donald (2001), humans are social creatures in a way that is different to the social lives of animals, and social interactions play a direct role in the development of the connections within the cortex of the human brain. Over time symbolic social interactions have allowed cultures to develop a stock of meanings and the motivation to 'tune in' to these meanings. Subtle emotional communication evidently plays an essential role in forming and sharing cultural knowledge. The

evolutionary development of symbolic communication appears to require a new coupling of emotions and their expressions to action, activity and movement in intersubjective encounters between children and adults.

1.2.2.2: Expression, movement and experience in infant learning

Gibson (1979) developed a theory that perception and action are reciprocally dependent. He suggested that awareness of one's self coexists with awareness of the environment and one's action within that environment. Lee (1998, p. 223) has demonstrated experimentally that prospective control of movements may be explained by regulation in the brain, for any form of movement, of 'time-to-closure of the motion-gap at the particular closure rate', a function defined mathematically by Tau Theory which gives precise formulation of Gibson's ideas about systems of action and perception. Movements that initiate actions and readjust them to perceptions of a continually changing array are not merely responsive to stimuli. They have to be planned before the action can take place; they have to occur in advance of perception.

Lee has developed and tested 'Tau' on a variety of movement tasks (2004). He has demonstrated that people control their movements in a predictable way that can be calculated by this mathematical procedure. Lee is currently working to provide more information about how emotional information is revealed in the 'pattern of flow' of expressive movements by variation of the dynamic energy that is required to alter amplitude in pitch, rhythm and timing and to change duration in intensity (2004). Lee's theory of how all actions are controlled is important for understanding communication because it addresses the essential questions about the expression of emotion in vocal and non-vocal gestures and how control of physical features of communicative signals can be a medium through which emotional meaning is expressed and perceived between human actors.

Rochat and Striano (2000) carried out research to observe how, through perception and action, an infant learns about its own movements and vocalisations. Through cross modal action and with self-exploration and the sensations that this creates, the infant can, they propose, *begin* to develop a sense of self. These authors found that by

2 months of age, infants were aware of their own proprioception or body movement sense. In particular infants were well aware of their own auditory experiences and, moreover, they seemed to be more responsive to interactions when their own movements and vocalisations were temporally matched with someone else's. This work leaves undecided if in fact early learning is necessary for infants to be aware of other persons' expressions as different from their own. Rochat and Striano assumed that the communicative function must be learned.

Darwin (1872), who developed the evolutionary theory of human emotions, stated that expression as shown through action of the face and body is "the language of the emotions" (p. 367). He observed three main principles of action and expression in humans and other animals. He believed that these principles could explain most of the actions and expressions that occur as a direct result of emotions and sensations in the mind. These principles were as follows:

1. Actions that are carried out as a result of a desire or sensation, if repeated over time, become routine and repeated automatically each time the desire or sensation is experienced. These actions are expressive in their nature, showing states of motivation of which the subject may be conscious.
2. As actions become habitual as described in the first principle, an unintentional propensity to act in the opposite direction will occur when an opposite state of mind is experienced. These actions may have no functional purpose, but they are extremely expressive.
3. The third principle proposes that actions can occur as a direct result of drives from the nervous system, without any influence of motivation or habit. These drives function in set pathways depending on which parts of the nervous system are activated and which parts of the body they are connected to.

Darwin suggested that actions and expressions can sometimes have elements of all of these principles and that they do reveal how people feel. He states that expressive actions that are seen in conjunction with certain states of mind have become innate

over long periods of evolution, although some of them still remain under the influence of learning. Darwin believed that many human expressive actions, including those made by infants, have developed from, or are utilised for, communication, and he published some of the earliest photographs of infants expressing emotions.

Darwin also noted the appearance of sympathetic expressions of emotion in his own son at the age of 6 months and concluded this was a demonstration of an 'instinct of sympathy' that when combined with expressive movements of the face and body provided a means by which preverbal communication can take place. Although Darwin's theory has been challenged many times over the last century by psychologists most concerned with the learning of emotional behaviours, his arguments about the innate foundations of emotions and their expression stand up in the light of modern studies, and are respected as providing an historical origin for systematic research into emotional expressions and actions (Dewey, 2001).

Crying and laughter are the earliest rhythmic vocal expressions of infants, and they are highly expressive. Even newborn babies make simple cooing vocalisations which may be precisely coordinated with similar sounds made by an adult imitating sensitively (Malloch, 1999). Babbling with articulation by movements of the lips and tongue, which begins about 6 months after birth, is recognised as the first rhythmic gesture that has a part in the development of speech (Masataka, 2003). Bates and Dick (2002) review evidence that gesture and rhythmic vocal activity are associated throughout the development of language and that both remain key elements in the expressive power of symbolic communication. This association is already obvious at the time babbling vocalisations develop. Masataka (2003) describes babbling and rhythmic hand banging as "sharing the properties of rhythmicity" (p. 168) and he suggests that hand banging is normally a developmental predictor of the onset of canonical babbling, a conclusion in line with other research that records coordination between these aspects of communication (Locke, 1994; Ejiri and Masataka, 2001; Iverson and Fagan, 2004).

Ejiri and Masataka (1999; 2001) found that during infant development, rhythmic motor activity and the range of frequency information in infants' vocalisations increased at the same age, and with new articulatory abilities, such as fast repetition of short syllables, and decreased after this level of skill had been acquired. They conclude that rhythmic action of the body and limbs and vocalisations develop together around the middle of the first year as a consequence of a centrally regulated transition in the development of cerebral systems for communication. Masataka (1995) found also that infants often extend their index finger when they were making syllabic sounds but not when they were producing sounds less related to speech. This research adds support to the idea that acoustic aspects of vocal communication are under the influence of cerebral regulators of rhythmic motor actions on many levels.

1.2.2.3 Emotions and experience in relating with other persons

Some categorisation of emotional experience is necessary in an exploration of how, and if, physical elements of body movement mediate mental aspects. Functionalist approaches to emotion have focused their debate around the categorisation of emotions as states excited in an *individual*, or they have considered the relationship between representationalist and non-representationalist explanations of socially important emotions categorised by cognitive psychologists as 'nonbasic'. Draghi-Lorenz et al (2001) review the theoretical distinction between the presumed 'basic' emotions (usually listed as *interest, joy, disgust, sadness, anger, surprise and fear*) which are taken as evident in the first 6 to 8 months of life, and 'nonbasic' emotions as those that are presumed to be only experienced when the individual has cognitive understanding of complex social situations and possible outcomes that can only occur in relation to 'another'. The required mental representations of other persons and their mind states are assumed to be acquired by social learning.

There are problems with any system intended to categorise elementary emotions which are evident when the natural variety and complexity of emotional expressions in everyday life are considered. There are many ways each supposedly 'basic emotion' can be experienced between persons. For example, a positive emotion could be categorised as simple *happiness*, but there will be other aspects of this emotion such as *positive interest, pride, joy, private pleasure, friendliness* or *affectionate*

sympathy. The natural 'ecology' of emotions is not so simple as psychologists have assumed, even in the world of an infant.

Draghi-Lorenz et al (2001) review theories on the development of emotion in infants and highlight the different theoretical perspectives concerning the role of higher cognitive abilities for the development of social (nonbasic) emotions. The main area of debate is over the issue of 'representation'. Is complex mental processing gained from experience essential for an infant's awareness and expression of emotions? The cognitive account proposes that even basic emotions can only occur after the age of 6 months, when it is estimated that emotional representations begin to form as a result of the infants learning from their own emotional behaviours and in social encounters. There is, however, a wealth of evidence to show that, on the one hand, infants below this age do have cognitive representations -- they form and organise representations and categories of fairly complex actions, causes, and attributions (Quinn, 2002; Friedman, 2002; Quinn and Schyns, 2003; Csibra, 2003; Durand et al, 2003) -- and, more pertinently, it has to be accepted that emotional expressions are innate regulators of communication from birth, and that these regulations have ancient evolutionary roots, as Darwin proposed (Panksepp and Bernatzky, 2002).

If emotions are an intrinsic part of the cerebral regulation of intersubjective awareness from birth, rather than an acquisition dependent on cognitive understanding of complex social relationships, and if this regulation is present as an innate feature of the human mind, then it is entirely possible that early forms and features of the 'nonbasic' emotions, some of which are highly effective in regulating even the simplest human encounters, will be also found from birth. Indeed, Draghi-Lorenz et al. say that, in their view, this would be the only solution for justifying a position which argues that nonbasic emotions do not rely upon a complex cognitive analysis of interpersonal encounters.

According to the nativist or Darwinian position, an infant is adapted to experience both complex emotions from birth: innately, the infant has the functional architecture that allows it to perceive social significance in the emotional expressions they

encounter when interacting with other persons. According to this view, infants are born with an innate capacity of relatedness with human others (Trevarthen, 2001). This means that they can connect with other sympathetic communicators and that they have the inbuilt motivation to do so. Likewise they will have the means to express their own affective state in communications that are mutually dynamic, contingent and reciprocal. Draghi-Lorenz et al quote Neisser as saying “what is perceived is not merely the other’s behaviour, but it’s reciprocity with one’s own” (Neisser, 1994, p. 400). If this is the case for an infant then we must conclude that the human brain and body are adapted with a set of physical channels that can be used to express, perceive and regulate the emotions that are necessary to regulate interactions.

There is further debate about whether the infant can be conscious of their emotions. Panksepp (2001) says “the experience of affect reflects a more ancient form of consciousness than that which sub-serves most of our cognitive abilities” (p. 14). Panksepp suggests that emotion and cognition are integrated processes although he stresses the point that they are still discrete and separately measurable functions in mental life and behavioural control.

Gergely and Watson (1999) assume that the infant has “little or no awareness of his or her dispositional states” (p. 101). However, infants behave in an organised, directed, purposeful way, especially in engagements with the coherent behaviours of other persons (Vandenberg, 1999; Fogel et al, 1999; Yale et al, 1999). Such purposeful integrity, and its communication, would be difficult to maintain (and perhaps without function) if infants were not prospectively aware, at some level, of their own and others’ motive states. Emotional experience can be considered as a connecting regulator between subconscious and purposeful processes or motives (Ochsner and Barrett, 2001) and functionally it would make no sense for the process and its regulation to be independent. By what means would an organism be able to perceive, express its emotions and co-ordinate its actions as one agent if it had no coherent awareness of these effects of its mind? It seems more likely that conscious perception is essential to inform purposeful processes and that cognition is a product of this connection, rather than its cause (Trevarthen and Reddy, 2007).

1.2.2.4 Emotional attachment, and development of feelings of the Self

Bowlby's theory of attachment conceives the infant as innately dependent on maternal responsiveness (Ainsworth and Bell, 1970; Bowlby, 1969; True et al, 2001). Beebe and Lachmann (2002) describe how attachment is built upon moment-to-moment synchronicity of parent and child interaction. They emphasise a shift away from the idea that this is a rigid or 'reflexive' instinctual process. They describe the optimal conditions for secure attachment as “a flexible balance of self and interactive regulation” (p. 105). In other words, the mother must not monitor or 'mirror' her infant too much, or too little, for their interaction to be mutually satisfactory. Jaffe et al (2001) found that this process depends upon optimal interpersonal timing patterns that are actively maintained by rhythmical communicative behaviours between infant and mother, and that it is affected by characteristics of both of them (Mangelsdorf et al., 2000).

Brain science attempts to explain how, by interacting with the environment, the infant develops a more 'explicit' or differentiated understanding of their own emotional feelings, at a higher 'cognitive' level. Interaction between developing core neural structures and the infant's affective environment work together to shape the way the individual will function emotionally. Pulvermüller (1999) describes a system whereby hearing, seeing, moving and feeling all combine to support the development of complex cell-assemblies in the brain. He suggests that with experience these assemblies form into established functional networks. He and others (e.g. Tzourio-Mazoyer et al., 2002) have found evidence that when infants look at their mother's face, areas of the brain that are later used for language are activated. There is also evidence to suggest that most of the full ranges of adult expressions of emotion are produced in the first months of life (Fogel et al, 2000; Sullivan and Lewis, 2003) which suggests that it is not the emotional states or their expressions but the infants' conscious awareness of them, and their uses in engagement with the world and other persons, that develops in the brain. As the infants' social world expands, the complexity of relationships increases, and with that, the ability of the child to explicitly 'label' or distinguish the feelings they have in ways that other persons comprehend and respond to.

1.2.3 Are Vowel Expressions linked to a General Intrinsic Motivation to Communicate?

This review has presented evidence that the development of communication in infancy is intrinsically linked with the development of other motor abilities. The question of how physical and motoric development is related to infants' ability to communicate is an interesting one. Information on how communicative structures develop in the infant brain gives some indication of how motivational factors of communication give communication behaviours symbolic meaning.

Tzourio-Mazoyer et al (2002) made PET scans on 6 infants at 2 months of age while the infants were looking at pictures of women's faces. They found that when infants viewed the faces, blood flow increased in a specific network of areas in their brains. Most interestingly, blood flow increased in the left superior temporal and inferior frontal gyri, areas of the cortex that are associated with language processing in adults. The authors suggest that the process of awareness so revealed links the infant's early preference for looking at faces with the later development of formal language. This research provides a clear indication of the importance of early experience of interaction with other persons' audible and visible expressions for later development of symbolic communication, but it does not identify the system or process in the brain that drives the infant to communicate.

A comprehensive review by Trevarthen and Aitken (1994; 2000) of brain development from before birth identified motivational systems of neurons that develop deep within the brain stem of the embryo during the first 2 months in the womb. They propose that early developments in the tissues of the brain stem fabricate an "Intrinsic Motive Formation" (IMF) (1994, p. 598). The IMF functions first to selectively configure brain networks that later respond to prenatal and then postnatal experience, including experience of other persons. It is proposed that it is the IMF that provides the infant with intersubjective abilities that regulate all future cognitive development. Trevarthen and Aitken emphasise that this development of a motivating system that directs brain development involves parts of the limbic system that are central in the expression and perception of emotion, and thus essential to the

establishment and regulation of communication after the infant is born. They note how abnormal development of brain systems during this time, before birth, can lead to communicative and emotional disorders such as autism.

Rizzolatti et al (2001) describe how areas of the brain associated with motivation and the recognition of motivated behaviours of other individuals are part of a functional brain network incorporating the orbitofrontal cortex and the amygdala in the limbic system, areas of the brain that are active in emotional control of social encounters. All motivating systems of the brain are 'environment expectant' - they seek particular engagements with the environment, and gain reinforcement in their development from engagements to which they are adapted (Bekoff and Fox, 1972). It has been shown that infants are highly sensitive to both the form and the temporal contingency of response from a caregiver, and that this responsiveness motivates infants to maintain involvement in interactions (Legerstee, 1997; Wijnroks, 1997; Stormark and Braarud, 2004). Contingency of an appropriate response from the human environment also seems to maintain the infant in a state of positive affect, suggesting that the intrinsic motivation to engage with others can be fostered (or disrupted) after birth, dependent on the infant's experiences in parenting care.

It is generally accepted that biological and emotional systems in the human brain are regulating behaviour from before birth (Bowlby, 1981; Schore, 2000). These theories offer an explanation of how psychobiological brain systems become transformed to generate and sustain psychological experience. There is substantial evidence that intrinsic motives allow an infant's psychological experience to grow through action and interaction within a human environment to which they are selectively responsive (Trevarthen and Aitken, 1994; 2000).

1.2.3.1 Purposeful and imitative vocal action and feelings

Vygotsky (1962) describes human speech as developing within “affective, expressive vocal reactions” (p. 40). However he states that although emotional aspects of this development are a means to make “psychological contact” he believes that they are “far removed from intentional, conscious attempts to inform or influence others”.

Modern research indicates that expressive vocal reactions of infants are actively utilised by infants as a way of regulating communications -- they do function as ways of informing and influencing others, and at some level they do so consciously and intentionally, though without any 'interpretation' or 'explanation'.

There are several important communicative behaviours that reveal that infants are more actively motivated and dynamic in their engagement with others than Vygotsky's comment would suggest. For example, when infants imitate the facial expressions of their caregivers during interactions (Bruner, 1996; Trevarthen, 2000), they are doing something that, while it appears to some degree to be automatic or 'reflex', provides a means by which infants can experiment with, 'understand' and learn how other people behave. Trevarthen and Aitken (1994) describe this process succinctly as "the correspondence between innate brain representations of the 'self' and 'other' as potentially equivalent" (p. 599). In this 'intersubjective' way, the infant begins to form more elaborate representations of both their own actions and the actions of others.

Smith and Bryson (1994) reviewed the literature for an imitative deficit in children from the age of 4 who have autism. They found that the evidence supported the idea that reduced imitative behaviour in autistic children is linked to a corresponding reduction in comprehension of any symbolic gestures. It is perhaps by the act of imitating in an intentional and adaptive way that the infant can gain the ability to understand symbolic gestures and in this process physical (acoustic) signals carried in other persons' "affective, expressive vocal reactions" are a medium through which emotional experience can grow and develop.

1.2.3.2: The role of vowels in the early development of communication

Although developmental linguistic research, intent upon understanding how reference and implication emerge, often neglects to consider the impact of emotional aspects of communication and speech, some interesting insights have been gained into factors that could allow infants to formulate predictions based on the expressive stream of acoustic stimuli that they hear in adults' speech. It is known, for example, that infants

can begin to differentiate their native language from other languages from about 5 months of age (Jusczyk, 1999; Nazzi et al, 2000). Long before they understand words, infants can discriminate the phonetic boundaries of words from languages with varying phonetic properties, and they can discriminate between different speakers in one language (Houston et al, 2000; Houston and Jusczyk, 2000).

It seems that rhythmic properties of speech help pre-linguistic infants to discriminate features corresponding to word boundaries from around the age of 7.5 months (Morgan, 1994; Jusczyk, 1999). Tincoff and Jusczyk (1999) played the words 'mommy' and 'daddy' to 6-month-old infants and found that they would turn towards a video of the corresponding parent. They did not find the same results when they used videos of unknown men and women. This study suggests that infants are not only able to discriminate boundaries but that they are able to understand the meaning of some commonly heard words. It is important that the words they understood in this experiment were *emotive* words in the sense that they named the infant's closest caregivers. However, the study in question does not address this issue, and it's difficult to know from linguistic research of this kind if the infant at this age has any measure of emotional or psychological states in people who speak. What is clear, however, is that infants can, around the middle of the first year, discriminate a complex range of phonotactic information, and it seems that discrimination of these features allows the infant to create some sense of order in the speech stream, so they can form expectancies about how their native language sounds, at least for often repeated simple utterances.

Linguistic research does appear to be consistent with other developmental research showing that infants are paying particular attention to prosodic and paralinguistic features of speech in specific emotional situations. Cheour et al (1997) used event related potential (ERP - a measure of brain activity) to explore infant brain activity when they presented different vowels to newborn and 3-month-old infants and found an increase in the ERP measurements for unfamiliar (manipulated) vowels, and a positive reaction to standard vowels. The response was stronger in the older infants. So very young infants can discriminate deviations from how vowels 'should' sound

and this ability seems to become more efficient over time. Fisher and Tokura (1996) carried out research to try and understand to what extent infants receive prosodic and pitch variations that they might discriminate. They analysed infant-directed speech from American English mothers and from Japanese mothers, recorded when the women were addressing their 13.5-14-month-old infants. They found that in both languages vowel sounds at the ends of phrases were lengthened. In the same year, Kuhl and Meltzoff (1996) found that when infants (aged 12, 16, and 20 weeks) were presented with adult vowel sounds, they often produced vocalisations that were imitations of the vowel sound presented. Kuhl et al (1997) studied the infant-directed speech from American, Swedish and Russian mothers. They found that when mothers addressed their infants, they lengthened their vowel sounds to give the characteristic pitch extensions of infant-directed speech. Kuhl et al (1997) suggest that this has the advantage of giving the infant information that highlights linguistic elements of the speech. Again, other intersubjective functions were not attended to. These findings encourage the conclusion that acoustic features of speech elements (especially vowel sounds) are worthy of further study to investigate the full range of functions that may be served by communication with infants.

Infant perceptual abilities and caregivers' infant-directed speech appear to be mutually adapted to enable the infant to have a much more complex awareness of acoustic parameters than was previously believed possible. The question here is how this complementarity can support the infant to be aware of and express its own affective and other psychological states. Leveille et al (2001) looked at the vocalisations of 41 mothers and infants and found that the majority of mothers' speech was related to how they perceived their infant to be *feeling*. Infants appear to smile more in response to their mother's smiling face (Messinger et al, 2001), and it is known that prosodic features of emotional infant-directed speech are discernable by exaggerated pitch shifts, such as those distinguishing different emotional expressions, rather than the prosodic features themselves (Trainor et al, 2000). So it seems that there is a strong link between emotional aspects of a mother's speech and acoustic features of her voice, and that this link, which is currently not completely understood, offers itself as an interesting candidate for further research.

1.3: The Research for This Thesis

1.3.1 The General Aims

To contribute to understanding of how infants develop from vocal communication of affective and motivational state to language within the first year of life, my research focuses on the following questions:

1. Is there a systematic relationship between acoustic features (pitch, intensity and duration) of mother and infant vowel sounds?
2. Are acoustic features (pitch, intensity and duration) of mother and infant vowel sounds varied consistently when mothers and infants are participating in ‘engaged’ and ‘disengaged’ naturally occurring communicative interactions?
3. Do Japanese and Scottish mothers use acoustic features of their vowels differently (according to conventions of the mature language)? Is there any difference in the way that mothers and infants in each country vary vowel length, pitch and intensity systematically to express emotional and motivational messages?
4. How do infants react to affective changes in their mother’s voice when mothers communicate specific emotions?
5. Do infant vowel-like sounds convey specific, clear emotional messages? Can others understand infant attempts to communicate specific emotions, including those serving complex interpersonal regulations?

1.3.2 The Specific Goals

Evidence has been reviewed to examine the role of vowels in mother-infant affective communication. Evidence suggests that infants pay attention to acoustic aspects of their mother’s voice and that these acoustic aspects can provide the infant with a way to make meaning and to reciprocate in emotional engagement. The evidence reviewed so far highlights the need for a theory linking acoustic features with affective communication and helps to provide a rationale for looking at vowel use in particular. To contribute to a clearer understanding of how acoustic features of communication

can mediate psychological experience, the more specific goals of the current research are as follows:

- Study 1 will provide detailed description of acoustic features of mother-infant communication and aims to define and describe vowel sounds produced by mothers and infants, relating this to theories, presented in the review, of the parameters that are important for communication. It will assess the development of this process in a sample of mothers and infants in the first year of life, in two very distinct cultural contexts.
- Studies 2 and 3 will explore the ways in which adults and infants understand the emotional content of vowel sounds.

1.4: Overview of the Thesis

Chapter 2 will focus the literature review further to explore the theory of Communicative Musicality (Malloch, 1997; 1999) and to assess how far this provides a clear framework for my research, which aims to describe vowels as a feature of communication that may link emotive state and Intersubjectivity, to more fully examine relevant features of vowel sounds, and to determine how acoustic features of interaction vary in the very distinct cultural environments of Scotland and Japan. Chapter 2 will also present the methodology and findings from Study 1. Chapter 3 will further review the evidence that justifies exploring vowel sounds in particular as a link between acoustic information and psychological and emotional experience. The Chapter will then report the rationale for and findings from Study 2 (The Emotional Voice Study), which explored whether the infants responded in consistent ways to manipulations of the emotions expressed in mothers' voices. Chapter 4 will review the evidence and report the findings from Study 3, which suggest that infants do express emotions in their vocalisations that can be consistently understood by others. Finally, Chapter 5 will summarise and discuss the principal findings and theoretical implications of the research and will address gaps in the findings.

Chapter 2: Study 1, Voices and emotions of mothers and infants in two cultures

2.1 Introduction: The Nature and Uses of Vowels

Chapter 2 will examine evidence that addresses the following research questions:

1. Is there a systematic relationship between acoustic features (pitch, intensity and duration) of mother and infant vowel sounds and how are these features related to the overall process of communication of emotion?
2. Are acoustic features (pitch, intensity and duration) of mother and infant vowel sounds varied consistently when mothers and infants are participating in 'engaged' and 'disengaged' naturally occurring communicative interactions? If so, does this suggest that mothers and infants use vowel sounds as a means to communicate and express their motivational and affective states?
3. Do Japanese and Scottish mothers use acoustic features of their vowels differently (according to conventions of the mature language)? Is there any difference in the way that mothers and infants in each country vary vowel length, pitch and intensity systematically to express emotional and motivational messages?

2.1.1. Vowels as a feature of communicative musicality that link emotion and Intersubjectivity

One challenge confronting theories of the development of vocal communication is that we are as yet uncertain how discrete features of the various sounds we hear in speech, such as the extended voicing of vowels, are linked in to the perception of specific motivational or emotional messages. Researchers in linguistics have for many decades tried to link micro features of phonology to mental representation of intention and understanding. Studdert-Kennedy (2002) describes vowels as 'points of contrast' and suggests that they have very close evolutionary links to the ability of primates to produce and imitate signals of facial and vocal expression.

Trehub and Nakata (2002) suggest that features such as vowel elongation are essentially musical in nature, and they propose that the acoustic features of communication, such as 'intensity', 'intonation contour' and 'rhythm', have an emotional foundation: "Music begins its life-long journey as an emotive signal with clear emotional consequences. From the earliest days of life, infants' waking hours are filled with music or music-like materials in the form of maternal vocalisations" (p. 37).

The Theory of Communicative Musicality (Malloch, 1999) proposes a strong link between the rhythmic and energetic regulation of all body movements and the rhythms and expressions of communication. There is evidence to suggest that emotional expression and communication depend upon the 'prospective' regulation of motor activity (Lee, 1998). Research with deaf mothers and infants has revealed common dimensions in different acts of communication; for example, rhythm is utilised for expressive manual activity in much the same way as it is for regulating acoustic expression of emotion in the voice (Masataka, 1992; Trehub and Nakata, 2002). Deaf mothers hand sign to their infants in a style that is as distinctive as 'infant directed speech'. Their sign language is more exaggerated and more repetitive. Trehub and Nakata (2002) describe this 'infant directed sign language' as "dance-like" (p. 38).

Mirror neuron theories have been mentioned, and they are relevant here for two reasons. First they provide strong physiological evidence for a link between action and perception in the brain, and second, they propose a mechanism that translates perception and action into a 'sympathetic' psychological experience between subjects. It is proposed that neurons fire in the monkey and human brains in response to perceiving the 'intentions' guiding actions and gestures of the hands and of vocalisations (Rizzolatti and Arbib, 1998; Rizzolatti et al, 2001; Rizzolatti and Craighero, 2004). When a person observes jaw movements that are salient in recognition of movements of the mouth that are used for speech, sympathetic activity occurs in areas of the observer's brain for perception and production of speech

(Shibukawa et al, 2003). It has been demonstrated that when an 8-week-old infant views even a stationary woman's face parts of the brain that regulate communicative expression in adults are excited (Tzourio-Mazoyer et al, 2002). It is possible that vowel sounds, and their intentional/affective messages, could be represented in a similar way in the brains of 4-month-old infants.

Evidence suggests that some sounds of other individual's actions cause 'mirror' neurons to fire in the brain of a listener. Kohler et al (2002) recorded neurons in the pre-motor cortex of monkeys. They found a group of neurons that fired in response to sight of actions that were related to specific sounds. These same 'multimodal mirror neurons' fired when the monkeys *saw* the action being performed and also when they *heard* the associated sounds, even if these were presented separately. It is interesting and relevant that these 'multimodal mirror neurons' are found in the pre-motor area of the monkey brain that is homologous with Broca's area in the human brain. Kohler et al state that "these neurons have the capacity to represent action contents: second they have the auditory access to these contents so characteristic of language" (p. 848).

This evidence adds weight to the idea that expressions of gesture and vocalisations are coupled, and shows that aspects of sound and movement are represented in the primate brain in a holistic 'psychological' way – aspects of visible action and sound are represented, or experienced, as one intention, in which each perceived event can be understood separately as indicating a particular action, each stimulus creating the same neuronal response on its own as they do together.

2.1.2. Prosodic phonology

This research will focus on vowel sounds as prominent components of speech communication and song that play a central role in the expression, perception and mutual regulation of emotion between mothers and infants. However, it is necessary to place extended vocalisations in relation to other expressions that are associated with them.

The theoretical framework presented in Chapter 1 emphasises the importance of the combined activity of vocal and non-vocal aspects of communication. The prosodic features of vocal communication itself are also complex, coordinating an array of separate elements, none of which might carry an emotive message of prosody on its own. A phonological approach to the study of prosody requires that features of pitch, intonation, loudness, tempo, rhythm and tone or quality of voice all have to be considered in combination when describing emotional expression in human voices (Crystal, 1997). These are features of a complex dynamic 'dance' of expressions that promote an understanding between communicating partners. It is important, therefore, not to focus on a single measured aspect of the acoustic signal to the exclusion of others, or indeed to the exclusion of accompanying non-vocal features of the communication.

Consonants, as articulatory modifiers of vowels, play a part in marking tempo and creating perceptible and meaningful units in the speech stream. They define syllables, and they give 'infant directed speech' a clear rhythm that helps the infant to become familiar with the language environment into which they are born (Ramus et al, 1999). Consonants may also be extended like vowels in powerful non-verbal expressions. For example, 'contoid consonants', such as /p/, /f/, /h/, and 'fricative consonants', such as s, can be prolonged to act in the same way as vowel sounds (e.g. pssst). They can be used to sooth (e.g. shhhh) and to express complex emotions (e.g. mmm). 'Vocoid consonants' such as l, r, w, or j can also act in the same way (Crystal 1997).

In the dynamic 'co-articulation' of vowels and consonants the speech elements are not presented as isolated events, one at a time, or in a discontinuous line. In reality these different types of sound “continually show the influence of their neighbours” (Crystal, 1997, p. 158). In other words, except when it is presented on its own as an exclamation, a vowel sound will always be contained by the consonant that came before it and it will colour the sound that comes after it.

Nevertheless, extended open voice sounds retain a salient capacity to convey emotions in all human communication. The speakers of African click languages, who

employ consonantal sounds in isolation as linguistic elements, also use vowels to express emotion (Takada, 2002). The Khoisan click language has a wider range of vowel sounds and more complex vowel system than most other languages, despite the fact that its special defining characteristic is the click sound (Crystal, 1997). So, while in a click language consonants are conspicuously used in a culture-specific way for some expressions of emotion, vowel sounds, similar to those of other languages, complement the expressive features of the click sounds.

It is concluded that a focused and more detailed analysis of vowel sounds within a wider description of all features of vocalisation and taking account of the syllables in which vowels are embedded, will bring reliable evidence as to which acoustic aspects become salient during the communication of emotion between mothers and young infants. We need a precise description of the various acoustic features of communication if we are to relate these features to psychological understanding of what feelings are being expressed.

2.1.3. The role of vowels in the vocal stream, and their function in communication with infants

It appears that parent-infant communication is greatly facilitated by the mutual co-ordination of rhythmical temporal patterning of vocal sounds. There are innate, communicative mechanisms, which facilitate the infant's awareness of, and interactions with, other people in phrased expressions of a mutually regulated 'narrative time' (Malloch, 1999; Trevarthen, 1999, 2001). These mechanisms coordinate motivational, sensory, motor and intersubjective systems and they operate within a cultural framework to which they are sensitive and from which they learn. It is possible that the salient acoustic features of infant and parent vowel sounds have evolved to highlight the rhythms of emotional narratives, and to coordinate temporal aspects of communication in ways that facilitate cultural learning.

Consonants mark the rhythm and phrasing of speech and, especially in a stress-timed language such as English, they define the tempo of speech (Crystal, 1997). The theoretical framework suggested in this thesis accepts the importance of consonants in

marking speech time, but postulates that vowel sounds may be considered to mark expressive temporal patterns at another level, providing information about when and how the beat and the rhythm is changing during a communicative interaction under the influence of emotions. The theory proposes that the length and quality of extended vowel sounds will assist the regulation of a melodic narration by which communicative partners transmit emotional and psychological information.

There are several reasons for believing that vowels carry important information about communicative rhythm in engagements with infants. They are among the first 'quasi-resonant' sounds to be produced by an infant (Oller, 1986; Fernald, 1989, 1993; Trevarthen, 1999; Locke, 1993). Vowels are salient features of all expressive adult speech, including, conspicuously, that used to communicate with infants (Ladd, 1996).

Vowels have more extensive harmonics (more 'resonant' formants) and infants show a poorer performance on perceptual tasks when harmonic information is limited (Clarkson et al, 1996). It is possible that infants may be drawn to vowel sounds because they are rich in perceptually attractive information and because they are placed in a way that highlights the phonologically stressed aspect of a language. Sensitivity to vowels correlates with the preference infants show for the affective content of infant-directed speech (Oller, 1986; Kitamura & Burnham, 1998).

Vowels play a dominant role in the experience of continuous vocal dynamics, allowing the listener to gain contextual information about the next sound to be made and its probable intonation (Ladd, 1996). Vowels, while they carry sequential and segmental details of speaking to the listener, also display socially important information about the gender, tone, psychological state and age of the speaker (Laver, 1980; Clarkson et al, 1996; Shimura et al, 1996). In this way they are relevant for affect and emotional expression in engagements and for the interpersonal narrative of any interaction, as well as the articulation of speech information. Vowel sounds can be conceived as extended and modulated vocal gestures. It seems likely that their modulations are features of the human respiratory tract that are adapted for signalling

motives, emotions and intentions, and that human hearing is specifically adapted to hear and interpret them.

2.1.4. Interpersonal 'contingency' in communication, and effects of its breakdown

It is widely accepted that emotional responses of a mother seeking communication with her infant must be directly 'contingent upon', or synchronised and emotionally sympathetic or 'attuned' with, the infant's actions, matching them in both timing and expressive qualities, if communication is to be satisfying (Murray, 1992; Murray et al, 1993; Robb, 1999; Stern, 1999). What is less clear is how infants detect the appropriate contingency of maternal expressions.

In Chapter 3 an experiment is described to test infants' awareness of the contingent quality of vowel sounds and of the emotional messages that may be contained in them. The experiment asks the following question: If emotional messages are distinguished and made predictable for infants by their perception of the quality of vowel sounds in mothers' voices when they are speaking to the infants, then how do infants react when only the vowel sounds of their mother's voice are changed so they are no longer appropriately contingent?

The importance of contingency of a subject's responses to stimuli is often explicitly expressed in justification of experimental research designs to test infants' prospective awareness and discrimination, and interactive contingency is manifestly of key importance in maintaining engagement between infants and caregivers (Yale et al, 1999; Stern, 2000; Jaffe et al, 2001; Reddy and Trevarthen, 2004). In a typically developing infant, a complex coordination of eye gaze, pointing, body movements, posture, facial actions and vocalisations is maintained when they communicate with an engaged communicative partner (Yale et al, 1999; Vandenberg, 2000; Fogel et al, 2000; Hsu and Fogel, 2001). This coordination is best maintained when the responses of a human partner are timed and modulated in relation to what the infant does -- when the intersubjective traffic of expressions is mutually regulated.

It has been shown that mothers who suffer from postnatal depression display communicative behaviour that is more likely to lack appropriate contingency with the interactions of their infants, and that this can have a long-term negative impact on cognitive and behavioural development for the child (Murray, 1992; Murray et al, 1993). Healthy mothers have been shown to produce imitative facial expressions in less than one second (i.e., one pulse) of their infant showing the same expression (Leveille et al, 2001). Mothers appear to be naturally attuned in to their infants' emotional signals and it seems that there are immediate links between maternal conscious and manifest awareness of their own emotional state and infants' expressive behaviour (Jaffe et al, 1999; Leveille et al, 2001).

Van Ergen et al (2001) studied the interactions between 150 4-month-old infants and their mothers, focusing on gaze, vocalisation and touch. They found that mutual responsiveness within interactions was organised around a three second time period (or a phrase length) and they suggest that mothers who were contingently responsive to their infants help them to maintain a shared attention focus with this temporal period or cycle. Tomasello et al (1993) interpret this intimate coordination as due to a combination of intersubjective imitation, active teaching and collaboration of learning between parent and infant. The evidence suggests that infants expect contingent or appropriately responsive behaviours from adults in mutually satisfying vocal and non-vocal communication.

Infants support the expressive intentions of mothers by imitating. Infants between 2 and 6 months old imitate vowel sounds in free vocal play with parents (Kugiumutzakis, 1993) and the emotions of 'interest' and 'pleasure' are conspicuous in these imitative exchanges (Kugiumutzakis et al, 2005). They are certainly capable of learning habitual patterns of exchange.

For the first 6 months of life an infant is exposed to repeating acoustic patterns in the maternal language in which vowels play a conspicuous part. However, there are many other expressive features of adults' communication with infants. If the infant does respond to differences in emotional expression of vowel sounds, this would not

necessarily mean that the critical emotional messages are contained only in those sounds, but would prove that infants can be aware of subtle differences in vocalisations that contain different vowels, and would suggest that vowels are playing an important role in communication with their mothers before the infants are attending to the messages of speech.

2.1.5. The culture of emotion and vocal expression

Communication of parents with infants by expression, perception and regulation of affective state occurs in a wider social context that influences the development of communication. Human communities are different in language and cultural traditions and different cultures foster differences in socially preferred expression of particular emotions. For example, Artic Utku people traditionally discourage their children from any outward expression of anger (Briggs, 1970).

Geertz (2000) says that we make sense of the world by constructing stories or narratives about everything we experience, and that these narratives are the expression of culture. He suggests that the mind and the culture it develops in are completely interwoven. Our brains or minds develop in a particular culture that then becomes a factor determining the way the mind thinks – mind and culture are context dependent, ‘complements’ of each other.

Fogel et al (2000) carried out research with 6- and 12-month-old infants to trace the development of smiling behaviour in game rituals. They found that at 6 months, infants showed appropriate differential smiling in response to different parts of tickling and peek-a-boo games. These games are played within a narrative framework in that they have a period of 'setting up' and a 'climax'. Infant smiles appear to express discrete aspects of enjoyment that were dependent on expected phases of the interpersonal context specific to each ritual of interaction. This demonstrates that emotional expression in infants is neither chaotic nor a product of immediate reflex reactions, but is part of a prospectively regulated expectation in dramatic narratives that have been learned.

Within cultures, individual family situations and the quality of self-other interactions have an important effect on the development of emotional processes and their communication. Neglected and physically abused preschool children showed differences in the ways in which they interpret emotional signals (Pollack et al, 2000). In comparison to control children, neglected children were found to have more problems distinguishing between specific emotional signals, and physically abused children showed a 'response bias for angry facial expressions' (Pollack et al, 2000, p. 679).

In a longitudinal study, Henderson et al (2001) found that maternal ratings of infant negativity at 9 months could be a factor predicting the child's wariness at 4 years of age. Henderson et al measured electroencephalogram (EEG) asymmetry and found that this relationship was only evident in male infants and children who had right frontal EEG irregularity. This team suggest that there may be some moderating role of sex on brain development reflected in the EEG cortical activity and affecting the development of social behaviour. Others have concluded that social factors are responsible for changing neurological and behavioural systems (Gorski, 2001). Panksepp (2001) states that negative emotions can restrict higher level processing and therefore place limitations on healthy emotional development. It seems that the innate physiological apparatus that generates and regulates emotions, and the environmental context with which it operates, will both affect development of infants' emotions.

Kahana-Kalman and Walker-Andrews (2001) used a preferential looking paradigm to examine infants' ability to recognise emotions in facial expressions. They found that infants aged 3.5 months could match happy or sad vocal expressions of their mother to her happy or sad facial expressions, but this discrimination was not evident when the infants were shown the facial expressions of strangers. This shows that intimate communication with someone close (in this case the mother) increases the infant's ability to effectively perceive emotion. Longitudinal work has also demonstrated the importance of social/contextual information in the early stages of emotional development and the long-term effect that disruption of relationships can have (Roe, 2001).

2.1.6. Vowel features, communication of emotion, and learning the 'mother tongue' -- how Japanese and English cultures differ

Single vowels produced by full-term (but not preterm) 6-month-old infants have been found to engage activity in a more extensive region of the vocal tract than multiple vowels. Infants use multiple vowel production to explore the movements of their tongue production as these are represented in the sounds produced: the first formant (F0) measures tongue height; the second formant measures tongue advancement (Goldfield, 2000). It may be concluded that single vowels should give clearer expression of the infant's emotion and this provides a rationale for using single infant vowels in the present study to compare communications in the two languages at a stage when emotional communication is of primary importance and the learning of multiple vowels in the mother's language has just begun.

English and Japanese languages have conspicuous differences in timing and syllable structure, and they differ significantly in the placing of vowels (Ladd, 1996). For example, in Japanese stress accent is indicated by pitch modulation only, whereas in English duration, intensity and vowel quality provide information about where the emphasis has been placed on a syllable (Beckman, 1996). It has been demonstrated that syntax in grammar influences infants' listening and voicing by 6 months, leading them to identify distinctive features of language and intention (Kuhl, 1998; Lecaneut et al, 1996; Trevarthen and Aitken, 1994). At 9 months, infants can show a preference for phonotactic structures that provide information on word boundaries even when prosodic features are filtered out, suggesting that they hear underlying structural features at this age (Oller 1986). It can be expected that the differences between Japanese and English vocal expressions will influence infants' vocal productions.

Japanese mothers, in addition to speaking differently, also adopt a different attitude to the motives and emotion of their infants from Anglo-Saxon mothers, giving greater value to interpersonal aspects, and less attention to their infant's interest in objects and events. They talk about different things with their infants, and their vocal expression is different (Shimura and Imaizumi, 1995; Bloom and Masataka, 1996). These differences may be related to the cultivation of manners in the two societies.

In Japanese society emotional cues are understood as central to communication and there are conscious differences in how you relate to other people, depending on how well you know them, their social position and whether the interaction takes place in public or private. The same could be said for Scottish society, but in Japan there are more clearly prescribed social constraints on how emotion is to be expressed in specific settings. This is demonstrated by traditional Japanese concepts such as *Koroko*. This is a Japanese philosophy to explain the integrated meaning of human nature. It values the combined action of heart and mind and a sense of knowing, which is different from the dualism of western thinking, which stresses the differences between 'dispassionate and objective' reasoning of the individual thinker and 'interpersonal and subjective' emotion. *Koroko* plays an important role in Japanese ideas about child rearing and has an effect on the ways in which emotional expression is communicated (Nakano, 1997).

For example, one of the most important virtues in Japanese society is to extend sympathy or 'kind consideration' (Nakano, 1997). Lewis (1995) found that an important part of elementary education is to “minimise competition and help children develop the feeling that we’re all in it together” (p. 7). The focus here is on what it means to be a kind, responsible member of the school community and this exemplifies aspects of *Koroko* that stress the importance of emotionally harmonising with others.

There are both similarities and differences in the ways in which mothers in different cultures respond to their infants. Bornstein et al (1992) found that culturally specific patterns of responsiveness were evident in mothers in France, Japan and the United States when they interacted with their 5-month-old infants. They also found that mothers in the three cultures showed a number of similarities in the way that they interact. For example, mothers in all three cultures encouraged their infants to explore the environment. All mothers imitated 'non-distressed' vocalisations produced by their infants, and they all behaved in a nurturing way when their infants were distressed. Differences lay in the ways in which mothers followed their infant's

gaze. American mothers responded more to their infants by directing their attention to the wider environment than the other mothers. Japanese mothers reacted more when their infants engaged in social looking and responded more by looking at their faces, than American or French mothers.

Despite the clear differences in cultural values and language, there are important similarities in Japanese and Scottish extension of pitch when talking to their infants. Masataka (1993) found that Japanese mothers used rising pitch contours when speaking to their infants and that there was a correlation between these contours and their communicative intention. It is not known, however, how vowels are utilised by mothers in the two cultures to express affect, or if they use them differently.

Anthropologists observe how culture can shape emotional expression throughout development. Weisner (2001) describes children as “apprentices to more experienced community members” (p. 1699). The large differences between Japanese childrearing practices and those that are common in Scotland can be expected to have effects even early in development. Suzuki (2000) reports that in Japan, infants and children are seen as being 'nearer to the gods'. They are therefore treated as being the most important members of the family, and perhaps of society, too. They sleep in their parents' bed until they are aged about seven, in the belief that this will strengthen emotional ties between the parents and the child. Suzuki describes this as a means to promote 'dependence', which is considered to be an advantageous characteristic in Japanese children. Infants in Scotland, as in other 'individualist' Western cultures (Bierhoff, 2002), are often encouraged to sleep in their own beds, and even in their own bedrooms, from the first year of their lives. They are expected to grow in independence and to acquire 'self-control'. It is unclear how these cultural differences in expectations about emotional dependence actually affect intersubjective communication in infancy – or if they do at all. However, it is necessary to keep such large differences in cultural expectations about emotional expression in mind, and attempt to control for them, when examining how mothers use their voice to regulate infant affective states, and this is the rationale behind my cross-cultural study.

2.1.7. Hypotheses of Study 1

Research Question (1):

Is there a systematic relationship between acoustic features (pitch, intensity and duration) of mother and infant vowel sounds?

It was predicted that pitch and intensity of vowel sounds would be coordinated with duration and that this would be evident in vowel sounds produced by mothers and their infants.

Research Question (2):

Are acoustic features (pitch, intensity and duration) of mother and infant vowel sounds varied consistently when mothers and infants are participating in ‘engaged’ and ‘disengaged’ naturally occurring communicative interactions?

It was predicted that mothers and infants would use acoustic features of their vowel sounds differently in ‘engaged’ and ‘disengaged’ periods of communication and that this would suggest that mothers and infants use vowel sounds as a means to communicate and express their motivational and affective states.

Research Question (3):

Do Japanese and Scottish mothers use acoustic features of their vowels differently (according to conventions of the mature language)? Is there any difference in the way that mothers and infants in each country vary vowel length, pitch and intensity systematically to express emotional and motivational messages?

It was predicted that Japanese and Scottish mothers would use acoustic features of their vowels differently but that both would consistently vary vowel pitch, intensity and duration throughout ‘engaged’ and ‘disengaged’ interactions.

2.2. Method

2.2.1. Participants and Procedure

To examine the above questions, 6 English-speaking and 6 Japanese-speaking mother-infant dyads (3 male and 3 female infants in each country) were recorded on digital TV and with digital audio in their own homes when the infants were 4 months old. Only mother-infant pairs were filmed to minimise confounding variables of voicing by different persons, including fathers, who might speak to infants.

Filming was arranged to suit the mother and her infant and so that at least 15 minutes uninterrupted and intimate communication was obtained while infants were alert. Mothers were asked to play with their infants as they normally do. Japanese infants were recruited in collaboration with Dr Koichi Negayama of Waseda University. Scottish mothers were recruited in and around Edinburgh.

Mothers were asked to complete a questionnaire to provide information about their nationality, their marital and employment status. Table 2.1 provides a summary of this information.

Table 2.1: Participant Information

<i>Infant ID Code</i> <i>S = Scotland</i> <i>J = Japan</i>	<i>Infant Gender</i>	<i>Mother's description of Infant Nationality</i>	<i>Mother's Age</i>	<i>Mother's Marital Status</i>	<i>Mother's Employment</i>
S1	Female	British	30	3	1*
S2	Female	British	38	2	1*
S3	Female	Scottish	40	2	2*
S4	Male	British	31	3	2*
S5	Male	Scottish	23	2	3
S6	Male	British	32	2	3
J1	Female	Japanese	33	3	1*
J2	Female	Japanese	29	3	3
J3	Female	Japanese	39	3	3
J4	Male	Japanese	32	3	3
J5	Male	Japanese	32	3	3
J6	Male	Japanese	36	3	3

Marital Status Code	Employment Status Code
1: Single and not living with partner	1: Full time worker *
2: Living with current partner	2: Part time worker *
3: Married and living with spouse	3: At home with child full time
	* = On maternity leave.

Mothers' ages were similar in the two countries: the mean age for Scottish mothers was 32.3 and 33.5 for Japanese mothers. Mothers in Japan (J) all described their nationality as being Japanese, whereas Scottish mothers either described their nationality as British or Scottish. As all the mothers were living in Scotland, for the purposes of this study they will be described as Scottish (S). All but one of the Japanese mothers was married and at home with their child full time. Most of the Scottish mothers live with their partners and had a full or part-time job. All mothers who worked were on maternity leave at the time of filming.

2.2.2. Analysis

The videos were viewed in their entirety and then a 30-second segment was chosen from the first ten minutes of each recording for each infant. The 30-second segment was chosen according to the following criteria for detailed micro-analysis of vowels, and for evaluation of the emotional engagement in the game:

1. The mother and infant were both participating in an interaction,
2. The infant was alert,
3. There was some rhythmical play by voice, gesture or physical contact between mother and infant.

As the filming sessions began all of the mothers addressed their infant to engage them in communication. Their interactions took the form of playful events on two main levels:

- a. Meta level *narrative* events expressing the overall purpose of the communication extending over tens of seconds or minutes,
- b. Micro level *phrases and expressive gestures* within the narrative shifting the initiative of communication from moment to moment.

This study chose to examine small games where one or the other of the interacting partners would initiate a communicative 'topic' or ritual episode, which was played out through repeating patterns of gesture and expression.

For example, the mother would begin to try and attract the infant's attention to her game by pushing the infant's feet in a rhythmical way and using her voice, often coordinating the rhythmical movement of her hands on the infant's feet, with a rhythmical sing-song aspect to her voice. These vocal and physical patterns would be repeated, sometimes increasing in tempo to maintain excitement and engagement.

Sometimes the games were initiated by the infant and the mother took up a rhythm from the infant. For example, the infant would show signs of avoiding an interaction by squirming, pulling away or turning their face away from the mother; this would often be accompanied by a complaining vocalisation (or cry) and an increase in the tempo or intensity of the infant's physical movements. The mother would then respond or 'attune' by imitating the altered tempo or intensity of the infant's expression. She would then repeat the pattern set up by the infant, either maintaining or increasing the tempo to continue interest and excitement, or slowly reducing the tempo to soothe the infant.

The beginning of a segment chosen for analysis would coincide with the beginning of one of the narrative events or 'games' defined as in the examples above.

2.2.3. Acoustic Data

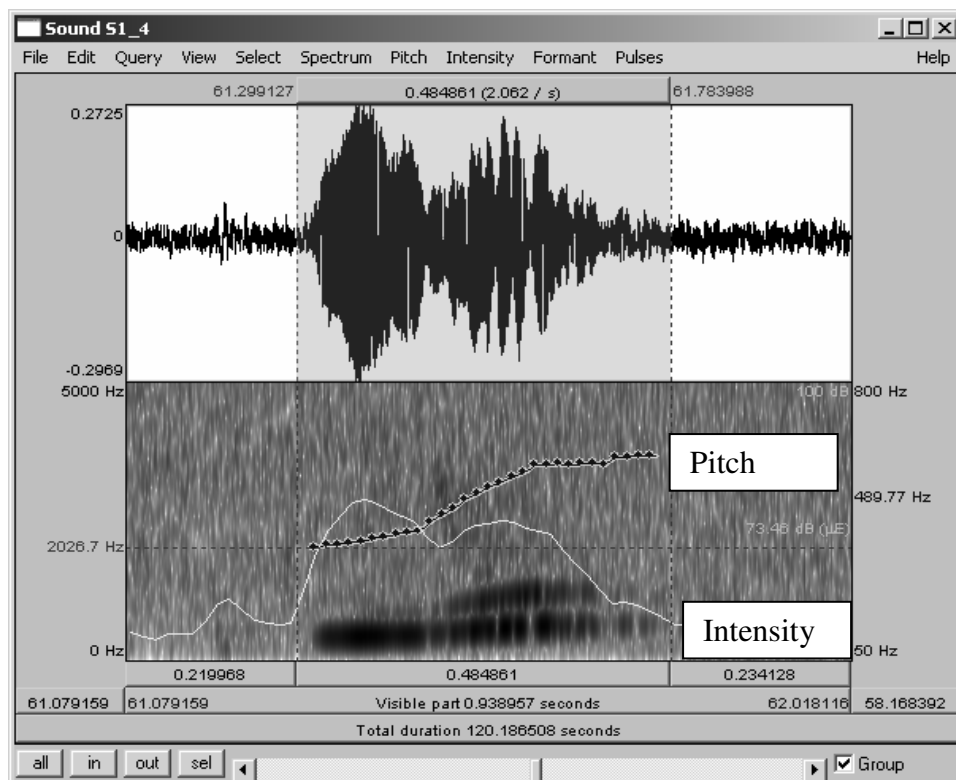
The 30-second audio segments chosen from the digital recordings for each mother infant pair were transformed into Waveform audio files using Sonic Foundry Sound Forge 6.0 professional digital audio computer software. Acoustic analysis was applied to the selected Waveform audio segments, using PRAAT computer software to measure pitch, intensity and duration for mother and infant vocalisations (1992-2005, Paul Boersma and David Weenink).

All vocalisations in each 30 second segment were analysed to locate vowel sounds in the naturally occurring speech for each of the mothers and the infants. Vowels were identified within the mothers' and infants' vocalisations using the following procedure:

- All vocalisations were listened to, to make a preliminary subjective identification of vowel sounds,
- Spectrograms for all vocalisations were viewed. Listening to the sound and viewing the PRAAT sound editor (shown in Picture 2.1) were the main identifiers used to make a subjective decision about where the vowel began and where it ended. The sound was listened to again to make sure that the vowel had been isolated from other elements in the vocal stream,
- Vowel *length* was measured in milliseconds.

The PRAAT program was then used to gain a precise measurement of *duration*, *pitch* and *intensity* for each of the identified vowel sounds. The program calculates the 'mean pitch' and 'mean intensity' for each vocalisation (in this case, each vowel sound) that is selected. Picture 2.1 provides an example and shows the measurement of pitch and intensity for a Scottish mother saying an 'oo' sound to her infant.

Picture 2.1: PRAAT Sound Editor Window



Some infants (2 Japanese infants, one male one female, and 1 Scottish female infant) did not vocalise during the segment chosen for analysis. Only the infants who produced vocalisations were included in the statistical analysis. If the mother and infant vocalised at the same time, separate pitch and intensity measures could not be calculated, and so pitch and intensity measures were presented for mother and infant combined. If pitch could not be calculated by the PRAAT program, as, for example, when the vowels were spoken too quietly, then the vocalisation was excluded from the acoustic analysis.

A vowel-length distribution analysis was carried out. All vowel sounds, by mothers and infants, were measured and categorised as follows:

- ‘Extended’ - over 250 milliseconds (ms),
- ‘Long’ - between 151 and 250 ms,
- ‘Short’ – between 50 and 150 ms.

It was predicted that the Japanese and Scottish subjects would use vowels differently and that this would be detected by differences in the mean pitch and intensity and by differences in the distribution of vowel lengths.

2.2.4. Analysis of Emotional Engagement

Using only the video data (frame-by-frame and without the sound), periods of ‘engaged’ and ‘disengaged’ communication within the segment was categorised from the visible gestures and expressions. The emotional quality of ‘engagement’ was judged by 2 Scottish raters, one of whom was blind to the purpose of the experiment. The blind rater was asked to identify periods of communication where the mother and infant were emotionally ‘engaged’ and ‘disengaged’ using an intuitive assessment. After a short period of practice both raters were able to independently and consistently identify ‘engaged’ and ‘disengaged’ periods of communication. 15% of all segments were rated by both raters with 100% concordance. The idea of ‘engagement’ is based on Stern’s (1985) theory of attunement. ‘Engagement’ was defined as established when mother and infant, *shared intimate moments of close*

mutual awareness of each other or of a shared interest. ‘Disengagement’ was established when the infant’s attention moved away from the mother or from the communicative interaction.

Statistical comparison, of acoustic features of mother and infant vowel sounds, in different emotional engagements aimed to answer the following questions:

- Are pitch, intensity and duration of vowel sounds, associated with expression of particular emotions in engagement?
- If so, is there any difference between Japanese and Scottish use of vowels?

If acoustic features of vowel sounds are associated with levels of emotional engagement, then it will be concluded that vowel sounds are a central feature used to convey emotion and establish emotional meaning.

2.2.5. Analysis of Rhythmic Tactile Play

Rhythmical play was not the main focus of this study but it was useful to explore this non-verbal aspect of communication. In particular it was important to explore how mothers used playful and sensitive touch to attract, excite and soothe their infants. Tactile play was categorised using the video data (without the sound), according to the following broad definitions:

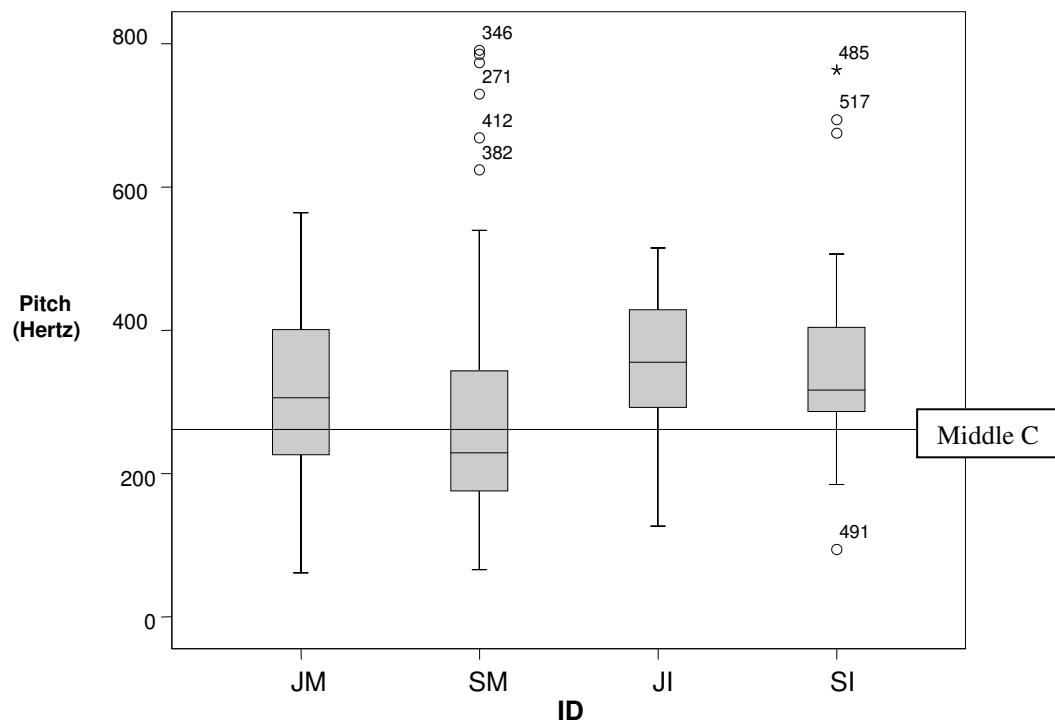
- ‘Active movement’: the mother swings the infant up, down or sideways,
- ‘Skin stimulation’: the mother strokes or rubs the infant,
- ‘Encouraging participation’: the mother holds her hands out to the infant, or helps the infant to make movements ,
- ‘Exercise play’: the mother exercises the infant’s limbs.

2.3. Results

2.3.1. Descriptive information

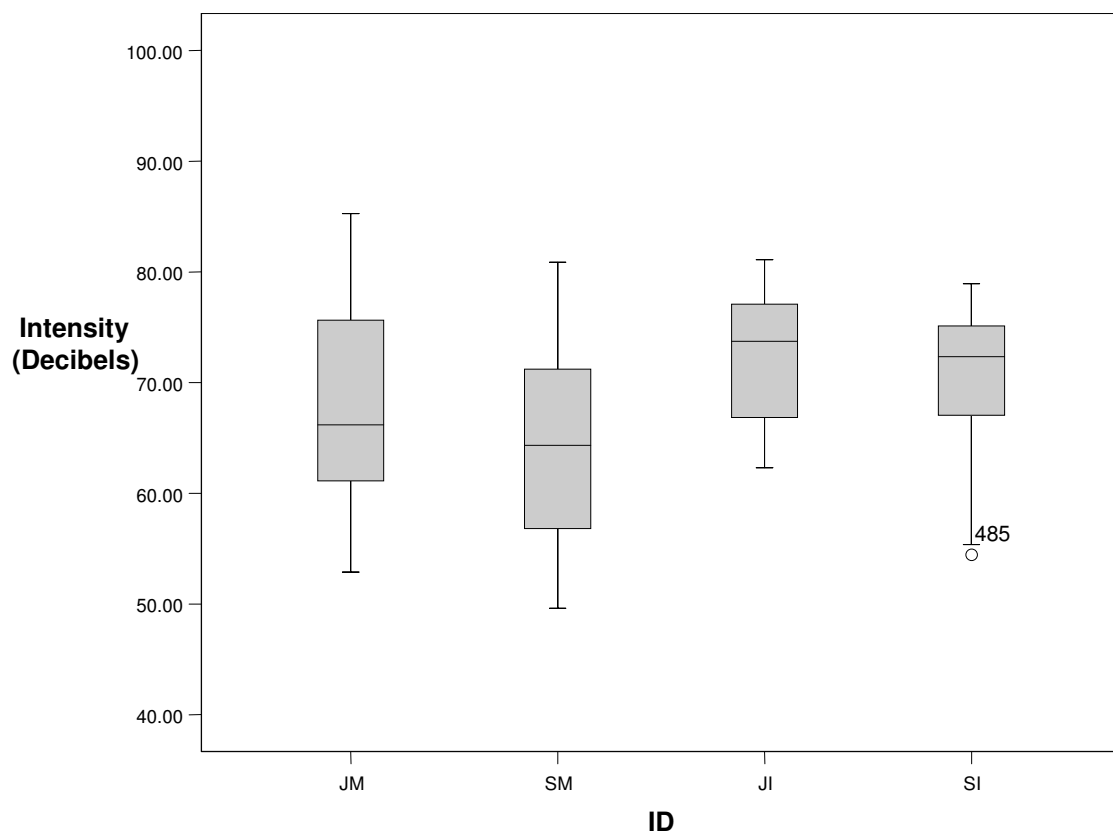
Graph 2.1 shows pitch information for vowel sounds produced by Japanese mothers (labeled JM on the graph), Japanese infants (JI) Scottish mothers (SM) and Scottish infants (SI). The graph shows that all of the infant and most of the mother vowel sounds have a pitch measurement above Middle C. The outliers show that some Scottish mothers and infants have used extremely high pitch during their interactions. Numbers marked on the outliers represent the case number of extreme vowels. The median of infants' pitch is higher than for their mothers and when mothers and infants are compared by nationality; Japanese mothers and infants have higher pitch than Scottish mothers and infants. Error bars show that there is fairly high individual variation in pitch for all groups. Distribution is similar in range for Japanese and Scottish mothers and is lower for the infant groups.

Graph 2.1: Pitch information for Japanese and Scottish Mothers and Infants



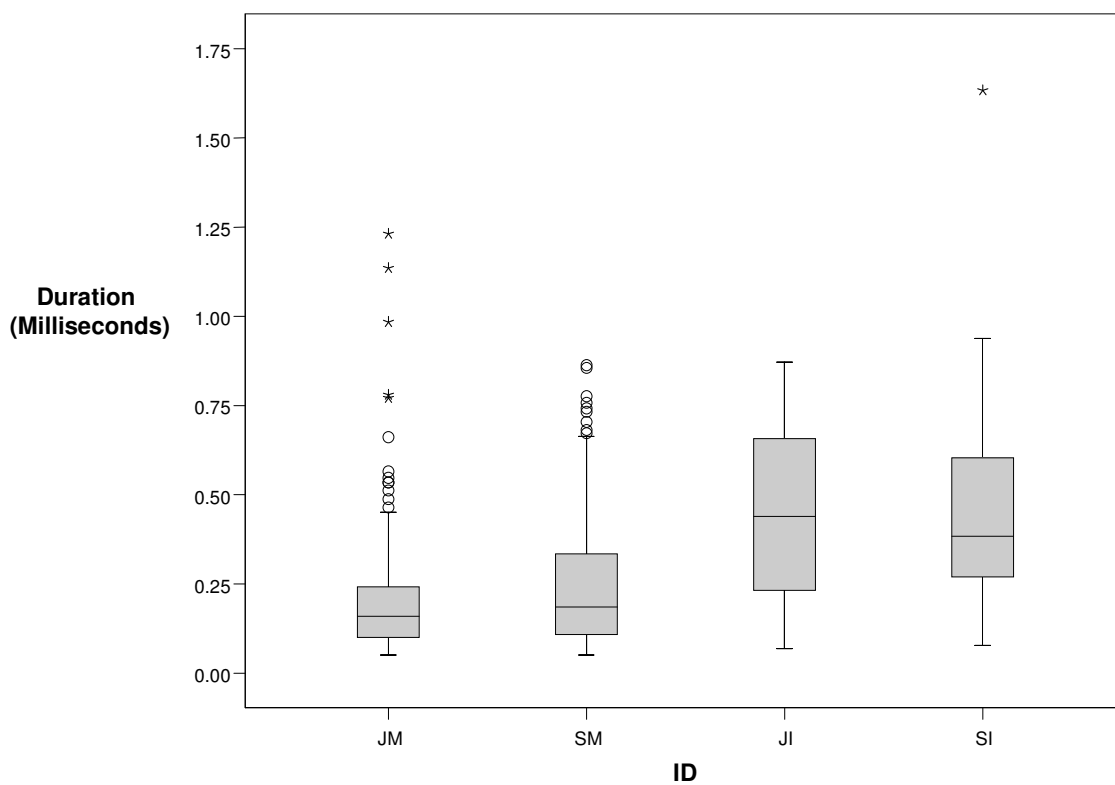
Graph 2.2 provides information about intensity of mother and infant vowel sounds in each country. It shows that infants appear to have higher intensity than their mothers but that they have less variation in the intensity of their vowel sounds than their mothers. There are fewer extreme cases for intensity than for pitch and error bars show that Japanese infants have less individual variation in their intensity than other groups.

Graph 2.2: Intensity information for Japanese and Scottish mothers and infants



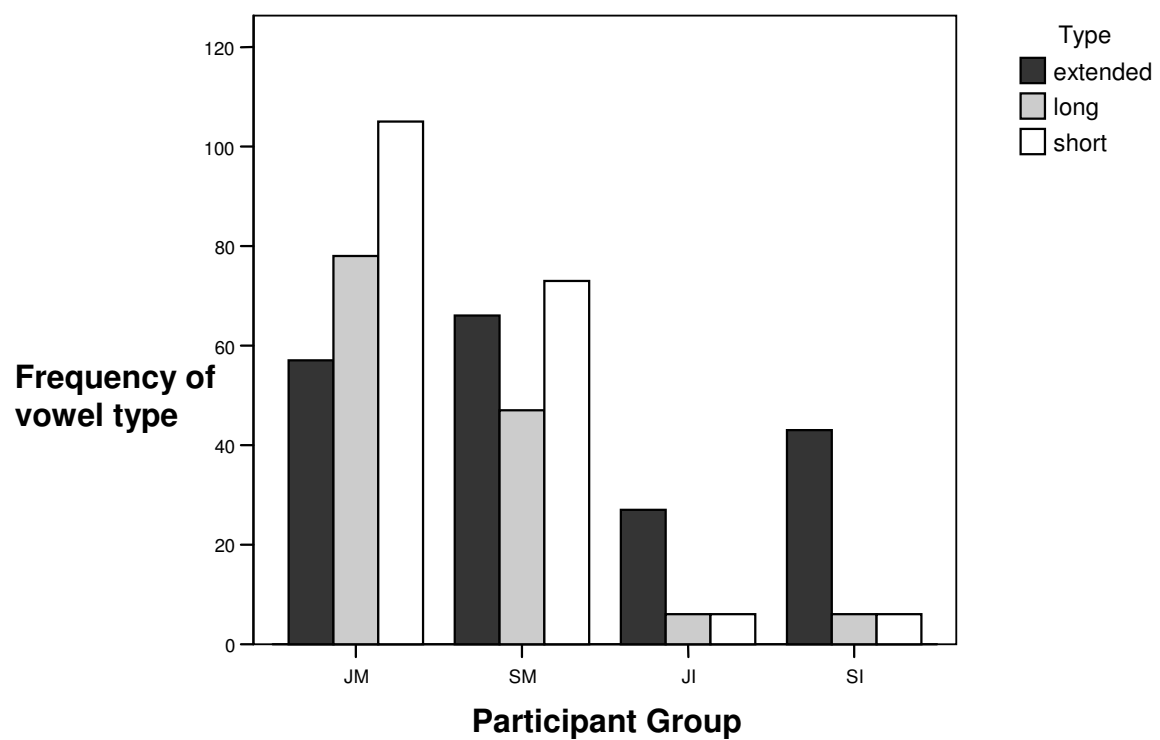
Graph 2.3 gives vowel duration information for Japanese and Scottish mothers and infants. It shows that mothers have more variation in the duration of their vowel sounds, some of them using extremely long vowels when interacting with their infants. Error bars show that there is more individual variation for vowels with longer duration than for shorter vowels. However the median of their duration suggests that they are using more ‘short’ vowels.

Graph 2.3 Vowel duration information for Japanese and Scottish mothers and infants



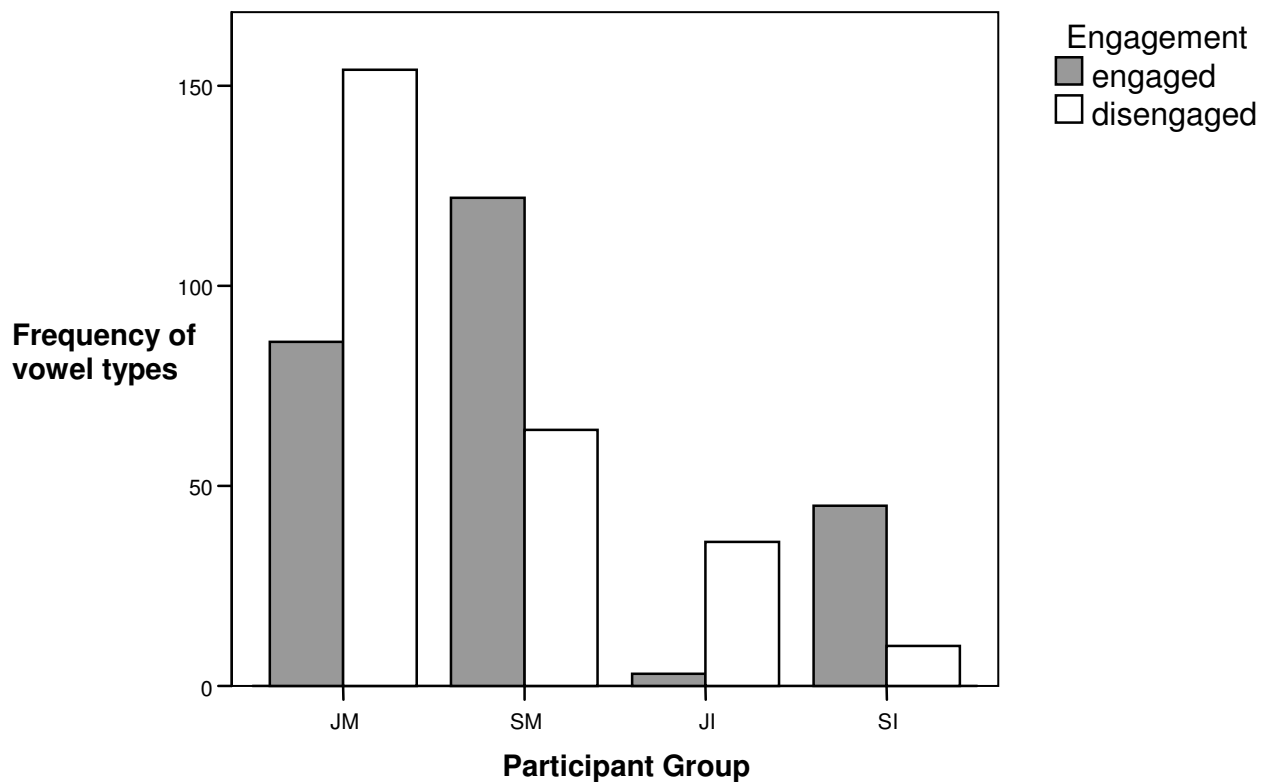
Graph 2.4 shows the frequency of ‘short’, ‘long’ and ‘extended’ vowels for Japanese and Scottish mothers and infants. It shows that Japanese mothers use more ‘short’ and ‘long’ vowels than Scottish mothers. Scottish mothers and infants use slightly more extended vowel sounds than Japanese mothers and infants and infants in general use more ‘extended’ vowel sounds than they do ‘long’ or ‘short’ vowel sounds.

Graph 2.4: Frequency of vowel types for Japanese and Scottish mothers and infants



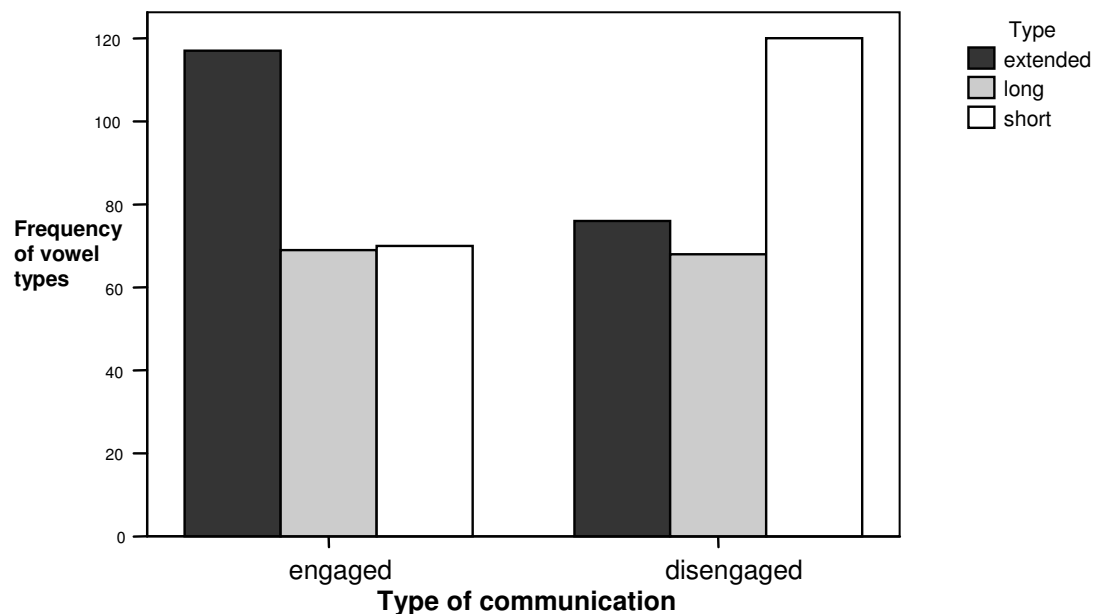
Graph 2.5 shows the number of vowel sounds produced by Japanese and Scottish mothers and infants during ‘engaged’ and ‘disengaged’ communication. Japanese mothers and infants produced more vowels than Scottish mothers during ‘disengaged’ interaction. Scottish infants produced more vowel sounds during ‘engaged’ interaction than Japanese infants, whereas Japanese infants produced more vowel sounds during ‘disengaged’ interaction than Scottish infants.

Graph 2.5: Frequency of vowels during ‘engaged’ and ‘disengaged’ communication



Graph 2.6 provides information about which type of vowel sounds mothers and infants used during ‘engaged’ and ‘disengaged’ communication. They use more ‘extended’ vowel sounds during ‘engaged’ communication and more ‘short’ vowel sounds during ‘disengaged’ communication.

Graph 2.6: Vowel type in ‘engaged’ and ‘disengaged’ communication



2.3.2. Acoustic features of vowel sounds

The figures suggest that there are individual differences in duration and pitch of vowel sounds, between mothers, between infants, between emotional situations and between cultures. A multivariate ANOVA was carried out on the whole population from both cultures and a significant difference was found between individual mothers for pitch, $F(11, 414) = 15.280$, $p < .0001$, intensity $F(11, 414) = 70.604$, $p < .0001$ and duration $F(11, 414) = 3.862$, $p < .0001$.

To explore these differences in more detail, acoustic aspects were compared for different vowel types. The means and standard deviations for different vowel types

are presented in Table 2.2, which shows that pitch is higher for ‘long’ vowels and intensity is highest for ‘extended’ vowels.

Table 2.2: Pitch and Intensity for Vowel Types

		<i>Type</i>	<i>Mean</i>	<i>Std. Deviation</i>
Mothers	Pitch	extended	305.8	136.0
		long	325.9	132.3
		short	267.3	109.9
	Intensity	extended	67.8	8.7
		long	67.4	8.5
		short	64.7	8.4
Infants	Pitch	extended	338.2	92.1
		long	422.2	183.4
		short	363.7	142.7
	Intensity	extended	72.8	5.1
		long	69.5	5.4
		short	62.9	4.4

A 3 (type) x 2 (acoustic parameter) multivariate ANOVA was carried out on mothers’ vowels. Type (‘short’, ‘long’ and ‘extended’) was the independent variable and pitch and intensity were the dependent variables. This analysis revealed a significant main effect for vowel type and pitch, $F(2, 423) = 8.711$, $p=.0001$ and for intensity, $F(2, 423) = 5.761$, $p=.003$.

A multivariate ANOVA analysis was carried out on infants’ vowel sounds and found that there was a difference for intensity (highest for ‘extended’ vowel sounds), $F(2, 91) = 20.418$, $p=.0001$ but not between the pitch used for each vowel type (highest for ‘long’ vowel sounds) $F(2, 91) = 2.858$, $p=.063$ (ns).

Mother and infant data were combined and a multivariate ANOVA was carried out to see if there was a relationship between acoustic features of mothers' and infants' vowel sounds. Table 2.3 provides descriptive statistics from the analysis. It shows that Japanese infants use higher pitch in their 'extended' vowel sounds than other participant groups, and that Scottish infants use a higher pitch in their 'long' and 'short' vowel sounds than other groups.

Table 2.3: Descriptive information for acoustic features of vowel types

		Extended		Long		Short	
		Pitch	Intensity	Pitch	Intensity	Pitch	Intensity
Japanese Mothers	Mean	328.8	69.9	344.7	69.4	277.4	65.4
	SD	128.4	9.0	113.6	8.1	101.8	7.9
Scottish Mothers	Mean	286.1	65.9	294.6	63.9	252.6	63.9
	SD	140.2	8.0	155.0	7.9	119.8	9.1
Japanese Infants	Mean	370.9	74.1	335.1	70.3	313.4	65.3
	SD	103.9	5.1	50.9	5.0	33.6	2.3
Scottish Infants	Mean	317.6	72.0	509.2	68.6	414.0	60.5
	SD	78.2	5.0	230.7	6.2	193.9	4.9

The multivariate analysis found that there was a significant relationship between pitch, $F(2, 514) = 9.322$, $p = .0001$, and intensity, $F(2, 514) = 13.254$, $p = .0001$, when mothers and infants used different vowel duration types.

Partial correlations (shown in Table 2.4) confirmed that there was a positive relationship between pitch and intensity (that they are coordinated) when duration of vowel sounds was controlled for, for mothers $P = .460$, $p < .0001$ and when mothers and infant data were combined values. These dimensions appear to characterise a single factor of 'effort' or 'emphasis' in mothers' vocal expressions. There was no positive relationship between the pitch, intensity and vowel length for infants.

Table 2.4: Partial Correlations between Pitch and Intensity

	<i>Mothers</i>	<i>Infants</i>	<i>Mothers-Infants</i>
	Pitch	Pitch	Pitch
Intensity	.460*	.006	.415*
	*p < .0001		

2.3.3. Emotional Engagement

To see if these relationships between pitch and intensity mapped onto the level of communicative expression, emotional ‘engagement’ was subject to multivariate analysis, with pitch, intensity and duration as the dependent variables and ‘engagement’ as the independent variable. Table 2.5 shows means and standard deviations for pitch, intensity and duration for ‘engaged’ and ‘disengaged’ interactions. For mothers, a significant main effect was found for intensity, which was higher in ‘engaged’ situations, $F(1, 424) = 14.763$, $p < .0001$, for duration, which was higher in ‘engaged’ situations, $F(1, 424) = 17.453$, $p < .0001$, but not for pitch, which was slightly higher for ‘disengaged’ situations. The same analysis was carried out on infants’ vowel sounds, confirming that there was no significant difference between acoustic features of infants’ vowel sounds in the different emotional contexts of the two levels of engagement, although trends were in the same direction as for mothers.

Table 2.5: Acoustic features of vowels in relation to emotional 'engagement'

	<i>Engagement</i>	<i>Mean</i>	<i>Std. Deviation</i>
Mothers	Pitch	Engaged	293.803
		Disengaged	297.336
	Intensity	Engaged	68.001
		Disengaged	64.845
	Duration	Engaged	.256
		Disengaged	.187

Infants	Pitch	Engaged	345.240	133.998
		Disengaged	359.400	94.778
	Intensity	Engaged	71.669	5.554
		Disengaged	70.546	6.497
	Duration	Engaged	.473	.292
		Disengaged	.421	.249

2.3.4. Culture

To see if acoustic aspects of vowel sounds differed in vocalisations produced by Japanese and Scottish mothers, a multivariate ANOVA was carried out with pitch, intensity and duration as the dependent variables and country (Japan and Scotland) as the between-subjects variable. Means and standard deviations are presented in Table 2.6. The analysis showed a main effect of country for pitch, which was higher in Japan, $F(1, 424) = 8.797$, $p=.003$, intensity, which was higher in Japan, $F(1, 424) = 14.616$, $p=.0001$, and duration, which was higher in Scotland, $F(1, 424) = 7.557$, $p=.006$. When the same analysis was carried out on infant vowel sounds, no significant differences were found for acoustic features in the two countries.

Table 2.6: Acoustic features of mothers' vowels in cultural contexts

		<i>Country</i>	<i>Mean</i>	<i>Std. Deviation</i>
Mothers	Pitch	Japan	311.514	116.100
		Scotland	275.090	137.138
	Intensity	Japan	67.768	8.491
		Scotland	64.603	8.457
	Duration	Japan	.201	.162
		Scotland	.247	.182
Infants	Pitch	Japan	356.575	91.570
		Scotland	349.045	131.439

	Intensity	Japan	72.155	5.676
		Scotland	70.385	6.211
	Duration	Japan	.449	.261
		Scotland	.447	.281

When mother and infant data were combined, a multivariate ANOVA showed that there was a significant relationship between acoustic features of mothers' and infants' vowel sounds. Pitch was significantly higher for Japanese mothers and infants $F(1,517) = 5.462, p = .020$ as was intensity $F(1, 517) = 28.640, p = .0001$.

To examine the relationship between country and emotional engagement, a chi square analysis was carried out which showed that Scottish mothers had significantly more 'engaged' interactions (66% as opposed to 34% for Japanese mothers) in the segments that were chosen, $X^2=37.138, df=1, p=.0001$. Scottish infants were also significantly more 'engaged' (82%) than Japanese infants (7.7%), $X^2=50.176, p=.0001$. Multivariate analysis was carried out on the mother and infant data combined and found that there was a significant relationship between the intensity (which was higher in vowel sounds produced during 'engaged' communication), $F(1, 517) = 32.842, p=.0001$ and duration (which was longer in vowel sounds produced during 'engaged' communication), $F(1, 517) = 7.714, p=.006$. This suggests that mothers and infants are coordinating acoustic aspects of their vowel sounds during emotional communication.

2.3.6. Rhythmic/tactile play

Acoustic aspects of mothers' vowel sounds were examined using a multivariate analysis to see if they were used differently in contrasting types of rhythmic/tactile play. Table 2.7 presents descriptive information for pitch, intensity and duration of mothers' vowel sounds during each type of play. The figures suggest that pitch is highest during play that involves the mother encouraging her infant to participate. The intensity and duration of mothers' vocalisations are highest during 'active movement' play where the mother is swinging, raising or lowering her infant.

Table 2.7: Descriptive information for acoustic measures of mothers' vowels during tactile play

		<i>Tactile</i>	<i>Mean</i>	<i>Std. Deviation</i>
Mothers	Pitch	active movement	295.248	106.122
		skin stimulation	248.190	112.266
		encourage participation	369.443	156.448
		exercise play	360.895	108.368
	Intensity	active movement	71.307	6.313
		skin stimulation	62.292	6.227
		encourage participation	69.448	6.507
		exercise play	71.206	9.963
	Duration	active movement	.265	.207
		skin stimulation	.221	.158
		encourage participation	.206	.163
		exercise play	.209	.186
Infants	Pitch	active movement	337.693	110.122
		skin stimulation	368.871	129.650
		encourage participation	384.913	90.235
		exercise play	326.683	37.487
	Intensity	active movement	73.214	4.234
		skin stimulation	69.542	6.117
		encourage participation	58.268	3.291
		exercise play	74.738	5.286
	Duration	active movement	.455	.240
		skin stimulation	.462	.318
		encourage participation	.229	.125
		exercise play	.449	.207

Multivariate analysis of mothers' vowel sounds confirmed that there was a significant difference for pitch, $F(3, 422) = 30.482$, $p < .0001$, and intensity, $F(3, 422) = 47.422$, $p < .0001$, but not for duration. For infants there was a significant difference in

intensity in different play situations -- it was higher in 'exercise' play, $F(3, 90) = 12.905$, $p < .0001$ -- but there was no significant difference for pitch and duration.

To see if there was a relationship between acoustic features of mother and infant voices, a multivariate analysis was carried out and showed that pitch was significantly higher for 'encouraging participation' play, $F(3, 516) = 22.414$, $p = .0001$, intensity was significantly higher during 'active movement' play, $F(3, 516) = 50.847$, $p = .001$, and duration was significantly longer during 'active movement' play, $F(3, 516) = 10.035$, $p = .0001$.

2.4. Discussion

Overall the results of the above analyses showed that there are relationships between acoustic features of mother and infant vowel sounds, and that these relationships occur on two levels – there are both large individual differences, and there are group differences. From the results presented here, it seems that individually mothers and infants create something like their own, unique private 'vocal culture'. They use acoustic features of vowel sounds in different ways depending on the particular game they are playing, their personality, their confidence, and on the way in which they respond to each other.

However, when all of the vocalisations were analysed, significant differences were found between pitch and intensity for vowel duration type for both mothers and infants. This positive relationship suggests that acoustic features of vowel sounds were coordinated in some systematic way by the motive of the subjects, and this was found when data for all mother and infant pairs was combined and explored.

There was a correlation between pitch and intensity when duration was controlled for, which suggests that there is some positive relationship between these acoustic features for all mothers. It makes intuitive sense that pitch and intensity would increase for vowel sounds of longer duration, but we should ask what this might contribute to communication. A strong relationship sustained between two acoustic features of

communicative expression would provide a way of manipulating the 'performance' of voice sounds to create consistent forms of psychological expression that might be readily understood as meaning something to others.

Many researchers have suggested that communication with infants is actively synchronised and regulated in quality of expression by both mother and infant (Stern, 1985; Papoušek and Papoušek, 1986; Trevarthen, 1999; Malloch, 1997; Gratier, 2006; Papoušek, 2007). Trevarthen (1982) suggested that features of vocal communication are adapted for semiotic use to give a specific meaning to the sounds (especially those that are produced by and for pre-verbal infants). The findings presented here would be consistent with this view.

However, there was no relationship between pitch and intensity and the length of infants' vowel sounds. Although pitch and intensity were used differently by infants for different types of vowel sounds, the differences were not significant. Iverson et al (2006) carried out research to explore the link between vocal and non-vocal behaviours. They suggest that infants have the physical apparatus to control movements and they have motivation to act, but it takes time to develop the skills to reliably carry out more complex intentional acts. It may be the same for motor control of voice production. At 4 months old, the infants participating in the present research, who are undergoing large developments in motor control of all parts of their bodies, are beginning to show signs of using features of their voices in coordinated and consistent ways, but at this stage they are only beginning to master the process.

Mothers' and infants' pitch and intensity were found to be significantly correlated. When mothers' and infants' pitch and intensity was explored together, findings suggested that mothers and infants were coordinating the acoustic features of their voices. Other research has found that mothers and infants reciprocally imitate many aspects of their pre-verbal communication and that this process of sympathetic matching changes with age (Papoušek and Papoušek, 1985; Feldstein et al., 1993; Jonsson et al., 2001; Jaffe et al., 2001). Pitch and intensity of mothers' vowels were found to be related to different vowel types, whereas this was not the case for infants.

So, the coordination found here must to some extent be attributed to the mothers responding to the infant vocalisations as if they were consistent. Previous research has shown that this is common; mothers respond to infants' communication, adjusting their response to support infants' abilities and so promoting an affective learning environment for shared meaning (Gratier, 2006; Papoušek, 2007).

Some have been led to believe that infants cannot produce or understand meaning in vocal communication. They explain coordination in preverbal vocalisations as a by-product produced, without consistent meaning for interpersonal communication, when infants are simply practicing control over their vocal and respiratory systems. However, findings presented later in this thesis suggest that infants can produce sounds that are consistently understood by others, and that they are sensitive to emotional messages in the vocal communication of their mothers. It is known that infants do pay close attention to vocal cues in communication (Parise et al., 2007), that they can discriminate subtle differences in vowel sounds (Kuhl, 1983) and that they can differentiate between different affective states according to their vocal dimensions (Nelson et al, 1999). Their ability to imitate adults voice sounds and to do so in emotionally regulated exchanges has already been mentioned (Kugiumutzakis, 1993; Kugiumutzakis et al., 2005).

The analysis carried out for this study aimed to explore this possibility in more depth. The results showed a significant relationship between acoustic features of vowel sounds for different levels of 'engagement' for mothers. Mothers' intensity and duration were higher for 'engaged' interactions with their infants. The concept of engagement here is based on Stern's (1985) theory of 'attunement'. Stern states that for attunement to occur, features of communication must be matched between mother and infant when they communicate. In his theory, attunement "is a performance of behaviours that express the quality of feeling of a *shared affect state*..." (p. 142, emphasis added). Two of the features that he lists as essential to attunement of expressions are intensity and duration, and this would correspond with the findings presented here and would support the idea that physical (acoustic) features of vocal

communication are a principal means by which emotional meanings are shared and experienced in a psychological way between mothers and young infants.

Although not at a significant level, pitch was slightly higher in 'disengaged' interactions, perhaps because the mother is trying to engage the infant by increasing the 'urgency' of her voice. The findings presented here provide evidence that acoustic aspects of vowel sounds in particular are used in a systematic way to express meaning to infants – in this case, *emotional meaning*. Perhaps intensity and duration are used together as a 'medium' by means of which a level of emotional meaning can be maintained in communicative contexts, whereas pitch is used to attract attention (engagement) and excite. The context examined here was rhythmical play, and mothers were all actively trying to engage their infants in close, intimate contact and they clearly used their voices to help the infants.

Although acoustic features of mothers' vowel sounds were different during the periods when their infants were 'engaged' with them, there was no significant difference between acoustic features of infant vocalisations. It may be that the mothers are working harder (vocally) than their infants are, to maintain the infants' attention. Perhaps the infant's vocal response to a period of intimate engagement with their mother is consistently regulated in some expressive way, but that this regulation is not reflected in pitch, intensity or duration of vowel sounds and so is not measured in the method used for this research. Hobson (2007) describes how engagement in communication is brought about by the infant's ability to identify with physical (bodily) demonstrations of emotion expressed by their communicative partners. If mothers vary their voices consistently between different emotional environments, as found here, this would support an infant's ability to construe meaning from vocal and non-vocal gestures, even though the infants are not producing differentiated emotional messages themselves.

Differences were found in pitch, intensity and duration between Japanese and Scottish mothers, suggesting that mothers are using their voices in culturally specific ways. It was interesting that, again, there were no significant differences found for the

vocalisations of infants. This corresponds with research that shows infants do not become sensitive to language specific sounds, or imitate them consistently, until they are between the ages of 6 to 12 months (Werker and Tees, 1984; Jusczyk and Luce, 1994; Jusczyk, 1997; Jusczyk, 1998; Nazzi et al, 2000; Dehaene-Lambertz et al, 2006; Tsao et al, 2006). In the study presented here, Japanese and Scottish mothers and infants were found to systematically coordinate acoustic features of their vowel sounds. So at 4 months old, infants from both cultures can participate in vocal exchanges and although they are not yet producing consistent changes in their voices, they are involved. Mothers seem to provide an acoustic framework for the infant to model.

Werker et al (2007) found that vowel sounds in speech directed to Japanese and English-speaking infants contain distributional cues that they suggest serve several purposes: to attract the infant's attention, to share emotional states and to support learning of the infant's native language. The current study found that mothers in both countries varied acoustic features of their vowel sounds differently. The results here suggest that although this process occurs in both languages, there are language-specific ways that mothers use acoustic cues.

There were differences in 'engagement' between the cultures. In this study Scottish mothers and infants had more 'engaged' interactions. This finding is limited by the fact that this 'engagement' was coded by Scottish coders for both the Scottish and the Japanese recordings. Furthermore, a particular type of communication was chosen for analysis. However, during 'engaged' and 'disengaged' communication mothers and infants coordinated acoustic aspects of their vowel sounds differently in Japan and Scotland, suggesting that vocal sharing of emotion is culturally specific. In future research it would be interesting to get Japanese people to code 'engagement' and 'disengagement' as there may be some cultural differences in perception of this element of communication, and it would be interesting to extend analysis to a range of communicative situations, both private, in the dyad, and public in the sense of being part of a larger social group.

Some researchers may be uncomfortable with the use of an intuitive or 'subjective' measure of 'engagement'. However, there is evidence to show that people can make consistent judgments about emotional states and interpersonal purposes, which are difficult to measure 'objectively'. Issartell et al (2004) found that participants could consistently agree when assessing emotional state. When they watched a performance and simultaneously made intuitive judgments, judgments were consistent for participants. Later when they were asked to say what emotion was present from memory, their judgments varied.

This is similar to the findings presented here, at least for Scottish assessors, because acoustic segments were consistently rated in the same way by two coders. Evidence presented in the Introduction and in later chapters also supports the idea that people readily and consistently understand emotions in others, even though their recollections or *post hoc* verbal accounts may vary greatly.

There were differences for pitch in different types of tactile play (which was higher in 'encourage participation' play) and for intensity (which was higher in 'active movement' play) but not for duration. It is interesting that pitch is higher when mothers are trying to 'encourage' or 'invite' the infant to take part. This corresponds with findings presented here on pitch-use in 'engaged' interaction. Previous research has found that rhythmicity of vocal and other motor systems is coupled and that the vocal and gestural systems develop in parallel (Iverson et al, 2006). Increased intensity in one system will lead to an increase of intensity of the other and conversely the same connection is preserved to express slower and calmer rhythms in communication.

The intensity of infants' vowel sounds was significantly higher when the mother was exercising her infants' limbs. Viewing the non-verbal data on video, the infants do not appear to be excited during this type of play. They are often lying on their backs and relaxed. However, they often seemed to want to change from this type of play to something more exciting and it may be that the increase in their intensity is to signal a

protest. Further analysis of non-verbal data would be required to come to a firm conclusion about this.

Overall, this study found that there was a systematic relationship between acoustic features of mothers' vowel sounds and that infants were beginning to show a similar pattern in their vowel sounds. Mothers and infants coordinated acoustic features of their vowel sounds and they did this differently (but consistently) during 'engaged' and 'disengaged' communication and depending on their culture. This does suggest that acoustic features of vowel sounds are an important medium through which emotional meaning is shared between mothers and their infants.

Chapter 3: Emotional Voices

3.1: Introduction: Mothers' Voice of Emotion and Infants' Listening

This section will review evidence that can help to address research question (4):

What role do vowel sounds play in the communication of a mother's affect in speech and song?

Evidence will be reviewed to examine the role of vowels in mother-infant affective communication within the theoretical background of Communicative Musicality. A review of linguistic and developmental theories should provide a rationale for looking at vowel use in particular.

3.1.1: Connections between physical aspects of communication and emotional experiences

Although this current research will not be studying the brain, it is relevant to review some of the evidence from brain research to understand more fully how behaviours of communication are translated into psychological experience. There is now a wealth of evidence on neuropsychological aspects of vocal and gestural communication. In particular there are theories of how emotional attachment to the mother regulates developments in the infant's brain.

Allan Schore (2000) defined attachment as an “interactive regulation of biological synchronicity between organisms” (p. 23). He states that the process is driven by affective motivations and he proposes that this is mediated principally by the orbitofrontal system and its cortical and sub cortical connections. Schore (2002) also offers some interesting ideas about *how* neurobiological aspects of development are synchronised with emotional sensitivity in infants and their carers. He emphasises that “the mother must be psycho-biologically attuned to the dynamic crescendos and decrescendos of the infant's bodily based internal states of autonomic arousal” (p. 13). He suggests that this process occurs during play as the communicative pair interacts,

and that the mother and infant become familiar with each others' rhythmic patterns. Because the orbitofrontal area is closely linked to the limbic system it provides a direct channel for regulation of emotional arousal, a link between emotional experience, motivation for action based on this experience and the autonomic nervous system, a system of interacting components that allows the brain to learn new patterns and adapt to new situations. Schore suggests that inner mental development is related to the physical shape or pattern that is found in physical stimuli, in that recurring patterns create expectancies for the infant. Subtle patterns or rhythms seem to provide the infant with the means by which they can begin to form representations of what the world is about. Schore points out the importance of close examination of the factors that regulate affective mechanisms.

Schore's review of affective regulations and the development of the infant brain concurs with earlier work by Stern et al. (1982) who used microanalysis of recorded play to examine the infant-directed speech that six mothers used in interactions with their infants, at the ages of 2, 4 and 6 months. They found that specific pitch contours were used in specific situations, such as when the mother was trying to maintain a positive affect in the infant, or to get the infant's attention. They state that only when "recurring patterns are recognisable can they become meaningful units of information" (p. 727). Stern et al. admit that their findings do not confirm that the infants will always respond to certain pitch contours in specific ways, and it is unclear as to how the infants experience or represent intonational patterns of their mother's vocalisation in their brains.

Donald (2001) points out that pure symbolic communication is already achievable by computational neural networks but that this creates a limited capacity for understanding. He states that biological (analogue) processes allow the development of more complex symbolic interaction. These combine motor and motivational systems in a hybrid circular process that directly affects the brain in the development of the individual, as it has in the course of brain evolution.

However, this does not explain what the underlying features of these processes are, and how they are dynamically coordinated during actual communication with human infants. It is not easy to reduce human communication to abstract ‘primitives’ at any level. Some may argue that it is not always useful to try and do so. To gain more understanding of the development of a pre-verbal vocal symbolic communicative system that leads to language, it is necessary to look at how a relatively small number of acoustic features can be combined to produce such rich, indeed limitless, dynamic capability.

From a linguistic perspective, Studdert-Kennedy (2002) points out the difficulty in finding underlying acoustic units that correspond to specific phonological features of speech. However, he makes the link between physical features of the exceedingly versatile human vocal apparatus and an increased ability in humans to produce phonologically contrasting sounds along ‘articulatory continua’, in other words to produce an infinite number of intended outcomes. Studdert-Kennedy says that vowels could be used creatively in this way, and he emphasises that their utilisation is influenced by cultural differences in languages.

3.1.2: Timing, prediction, anticipation and narrative

If vowels do provide a link between physical behaviours and experiential processes, they will be utilised in the formation of intentional narratives that give an account of the experiences and actions of protagonists. Narratives are most commonly known as ‘stories’ with referential content but they can take other phatic forms. Narratives can be defined as events that occur in a purposeful sequence, which are related to each other by their generation, and which are meaningful. Infants may not understand the semantic content of a story, but they may still enjoy an emotional narrative that takes them through familiar and expected series of sounds, organised in familiar patterns of ‘set up’, ‘excitation’, climax and ‘resolution’.

The present research presupposes that infants are born with the ability to understand (and perceive) narratives in human vocalisation at a basic, emotional level, and that this helps them to predict and anticipate acoustic patterns in their mothers’ speech or

song. Infants appear to look for narrative sequences, or at least pay attention to them on a variety of levels.

For the purpose of this research, narratives are defined as vocal or non-vocal events or behaviours that are sequential, and are linked to each other in some purposeful and predictable way. These events can be linked in several ways, such as semantically, emotionally or rhythmically. A narrative will have a beginning and an end and may occur on several levels. For example, a game played by a mother and her infant can be considered as a narrative. There can be an understanding between an infant and their mother of the focus or 'aim' of the game and their actions and interactions will be related and have contextual meaning, in terms of the particular game that they are involved in. Pre-linguistic infants can join in with a narrative expressed in a familiar song (Powers, 2000) and it seems likely that they might use their inbuilt ability to perceive vocal or non-vocal events as being related, to help them anticipate or predict what will happen next – in other words to help them make sense of significant events in the world around them.

One factor to be considered is how discrete phonological or prosodic features are patterned in repetitive sequences. How they are rhythmical in their nature. Rhythms help to hold interactions within a time frame: they relate isolated acoustic aspects together in phrases and they provide a framework for variation and diversity in the succession of phrases. Intonation in infant-directed speech has very particular rhythmic characteristics and this seems to be the case over many language groups (Grieser and Kuhl, 1988; Papoušek and Papoušek, 1991; Fernald, 1992; Papoušek, 1996).

Van Egeren et al (2001) examined interactions between 150 mother-infant pairs. They found that much communicative and expressive behaviour showed temporal regularity around a 3 second interval. These phrases of behaviour were mutually contingent, and incorporated, in the course of play, expression of individual mother and infant differences in temperament. They were also contextually dependent,

suggesting that communicative development cannot be separated from situations capable of affecting socio-cultural development.

3.1.3: Infants' perception of speech and characteristics of infant-directed speech

Although linguistic research often neglects emotional aspects of communication and speech, it has provided some important insights into factors that could allow infants to formulate predictions based on the stream of acoustic stimuli that they hear in speech. It is known, for example, that infants can begin to differentiate their native language from other languages from about 5 months of age (Jusczyk, 1999; Nazzi et al, 2000), whilst also being able to discriminate words in a language with varying phonetic properties and between words uttered by different speakers (Houston et al, 2000; Houston and Jusczyk, 2000).

It seems that rhythmic properties of speech can guide pre-linguistic infants to discriminate word boundaries from around the age of 7.5 months (Morgan, 1994). When Tincoff and Jusczyk (1999) played the words 'mommy' and 'daddy' to 6-month-old infants, they found that they would turn towards a video of the corresponding parent. They did not find the same results when they used videos of unknown men and women, suggesting that infants are not only able to discriminate the word boundaries, but that they are able to understand the meaning of the words as names for particular persons. It is interesting that the words they understood in this experiment were emotive words in the sense that they named the infant's closest caregivers. However, the study did not address this issue directly, and is difficult to reach a firm conclusion about the infants' abilities to translate emotional meaning from physical features of verbal communication. This highlights an interesting gap in research findings, which provides a rationale for further exploration of this idea. What is clear, however, is that infants can discriminate a complex range of phonotactic information, and it seems that discrimination of these features allows the infant to create some sense of order out of the speech stream, so they can form expectancies about how their native language sounds.

Infant perceptual abilities and caregivers' infant-directed speech appear to be complementary in such a way as to enable the infant to have a much more complex awareness of acoustic parameters than was previously believed possible. The question here is how this combination can support the infant to be aware of and express its own affective and psychological states. Leveille et al (2001) looked at the vocalisations of 41 mothers and infants and found that the majority of mothers' speech was related to how they perceived their infant to be feeling, which they judged from the infant's facial expressions. Mothers adapted their speech in response to the perceived emotions of their infant.

Trainor et al (2000) examined prosodic features of infant-directed speech and found that mothers express heightened emotion in their voices when they talk to their infants. When expressing 'love-comfort', 'fear' and 'surprise' mothers used *pitch*, *intonation*, *tempo* and *rhythm* in different ways to express each emotion. This research did not examine the infants' psychological experience in response to their mothers' emotional voices, but it seems that there is a strong link between emotional aspects of a mother's speech and acoustic features of her voice, a link that is currently not completely understood, and which offers itself as an interesting candidate for further research.

3.1.4: Vowels as a link between acoustic information and psychological and emotional experience

There is some evidence to suggest that acoustic aspects of vocal communication can provide infants with a means of expressing psychological and emotional state. Papoušek (1992) found that some acoustic features of infants' vocalisations, such as duration and fundamental frequency were correlated with specific state-related information. She comments that this is evidence that "physical structure of pre-syllabic interactional vocalisations reflect the infants' behavioural-emotional states" (p. 238). Papoušek carried out further studies presenting isolated infant vocalisations to first time parents, multiparous parents, speech therapists and children. She found that although all groups could reliably detect cries and sounds of discomfort, only mothers and fathers who had infants the same age as those who had produced the test vocalisations could reliably identify 'joy' sounds. She replicated these findings with Chinese and American subject groups. Papoušek concludes that isolated aspects of

infant vocalisations provide discrete affect information. Interestingly, all groups tended to relate each vocalisation to the one they had heard previously. This suggests that although listeners can detect affective information from isolated vocalisations they are still trying to identify affective state in an overall context of ongoing, purposeful activity.

Others have had similar results with vowel sounds in particular. Shimura and Imaizumi (1996) coded the emotional content of isolated infant vowel sounds (from infants aged 6, 9, 12 and 17 months) and then presented them to different groups of listeners (mothers, students and nursery teachers). They found that correct judgement by the listeners of the emotional content of the sound (i.e. judgement that matched that of the researchers) was influenced by several factors. The listener group had an effect on listeners' ability to make correct judgements. Mothers were more sensitive to 'demanding/angry/rejecting' emotions than other groups. The emotional content of the infant vocalisation also affected judgements. Negative emotions were more likely to be correctly identified by all listener groups. Infant age also played a role in correct identification of the emotional content. Adult participants' ability to judge emotion increased as infants got older. So, even with no contextual information, listeners can perceive emotional content in isolated infant vocalisations – infants can produce vocalisations that provide discrete information about how they feel.

The theory of Communicative Musicality (Malloch, 1997, 1999; Trevarthen, 1999) and that of Coordinated Interpersonal Timing (Jaffe et al, 2001) would suggest that it is more relevant to look at how acoustic features of communication are linked to motivational aspects of interaction, and especially their timing. Trevarthen (1999) describes how "shared participation in the expressive phrases and emotional transformations of vocal games can facilitate not only imitation of speech, but interest in all shared meanings, or conventional uses, of objects and actions." (p. 155). The theory of 'Communicative Musicality' emphasises motivational aspects of communication and describes acoustic features of voices (such as *intensity*, *pitch*, *timing* and *timbre*) that are linked specifically to the quality of expression, perception and psychological experience of emotion. Ross et al (2007) suggest that humans

have a predilection for sounds that occur in musical scales and that this “arises from the routine experience of (musical) intervals during social communication by speech” (p. 9854). Miall and Dissanayake (2002) analysed infant-directed speech and found that mothers structure their interactions using features of poetry, such as 'metrics', 'phonetics' and 'emphasis'. In particular Miall and Dissanayake analysed positioning of vowels and consonants to give a measure of ‘presence’. They found that ‘presence’ was a clear indicator of intimacy (“mutual attentiveness”) in the communication.

Others have found that there are some patterns of acoustic information that are consistently associated with specific emotional states. Scherer (1992) provides a review of research into non-linguistic vocal affect signalling in adult humans and in animals. He reports a positive relationship between acoustic features such as *fundamental frequency*, *intensity* and *intonation*, and specific emotional states. Ladd et al (1984) found a clear relationship between *fundamental frequency range*, *voice quality* and *intonation contour* and adult listeners’ abilities to judge speaker emotion.

In the present research the main focus will be to look at how acoustic features of vowel sounds are consistently regulated in mothers’ infant-directed speech to express specific emotions. It will explore how infants respond to changes in the mother’s voice. In order to begin this exploration it is useful to focus on isolated physical or acoustic features in an effort to determine at what level perceptual discrimination of affect can take place.

3.1.5: Infants’ perception of vowel sounds – what features are salient and how are these features related to expression of affective state?

Research investigating the development of infants’ developing sensitivity to linguistic features suggests that infants are paying attention primarily to prosodic and paralinguistic features of vowel sounds in specific emotional situations. Cheour et al (1997) presented different vowels to newborn and 3-month-old infants and found an increase in the negative ERP measurements for unfamiliar vowels and a positive reaction to standard vowels. The reaction was stronger in the older infants so Cheour

et al conclude that even very young infants can discriminate deviations from how things 'should' sound and this ability seems to become more efficient over time. Fisher and Tokura (1996) carried out research to try and understand how infants use prosodic and pitch variation as cues to identify language specific prosody. They analysed infant-directed speech from American English mothers and from Japanese mothers, when they were addressing their 13.5-14-month-old infants. They found that vowel sounds at the ends of phrases were lengthened by mothers in both language groups. They conjecture from this that infants would be able to use vowel lengthening as a guide both for discriminating their native language and for providing a structural representation to aid in their own production.

As already reported, in the same year, Kuhl and Meltzoff (1996) found that when infants were presented with adult vowel sounds, they often produced vocalisations that were imitations of the adult vowel sound presented. Kuhl et al (1997) studied the infant-directed speech from American, Swedish and Russian mothers. They found that when mothers addressed their infants, they lengthened their vowel sounds to give the characteristic pitch extensions of infant-directed speech. Kuhl et al (1997) suggest that this has the advantage of giving the infant information that highlights linguistic elements of the speech. These findings provide further suggestion that acoustic features of speech, and in particular vowel sounds, are worthy of further study to understand whether they provide a medium through which emotional meaning can be shared between parents and their infants.

If vowel sounds carry important information to enable infants understand emotional messages, then one prerequisite would be that they are among the sounds that are easy for infants to perceive. There is evidence that infants can perceive distinctive features of their mother's voice in the womb, and indeed they show a learned preference for her voice when they are born (Fifer and Moon, 1995). Yamanouchi et al (1990) used the stomach as a model for the uterus, to test foetal discrimination of vowel sounds in particular. The subjects swallowed a microphone that recorded what could be heard in the stomach. They used Japanese vowel sounds for their stimuli and found that these sounds were clearly discriminable by the internal microphone. Infants' auditory

systems are known to be functional by the last trimester of gestation (Fifer and Moon, 1995) and it seems likely that by the time an infant is born, they will already be paying particular attention to vowel sounds.

Research by Kuhl (1998, 2001) has led to some interesting findings on the perception and production of vowel sounds by infants. She found that infants can distinguish vowel sounds produced by different speakers and concluded that infants must have the ability to classify equivalent sounds and that this must be a prerequisite for vocal imitation. However, infant vocal apparatus is not sufficiently developed to produce an identical match for absolute formant frequencies of vowel sounds. Kuhl concludes that infants may be able to create a representation of the sound they hear that is not reliant on specific frequency information, of sounds which they cannot produce and may not yet be able to detect. In other words, infants do not have to match formant information to know that the vowel sound they hear is the same as another vowel sound.

Kuhl's work (1998, 2001) also revealed that infants could detect phonetic changes in vowel sounds even when provided with irrelevant pitch contour change. This suggests that subtle changes in vowel sounds are noticeable to infants, even if other paralinguistic information is manipulated. In Kuhl's experiment, infants habituated more quickly to vowel dimensions than they did to dimensions of pitch alone. Kuhl concluded that infants' heightened ability to discriminate features of different vowel sounds without perceiving many detailed acoustic features was a means to reduce memory load.

Others agree that those acoustic features that they experience as easily discriminable will help to support infants to process and memorise phrases and words that are significant for them (Trehub and Nakata, 2002). If this is the case, then on another level it is possible that if vowels are easily discriminable they could also assist the pre-linguistic infant to segment speech and memorise syllable structures and rhythm classes that are common to the language that the infant hears spoken every day.

Phoneticians have found that it is possible to model this process of vowel perception with neural networks. Boersma et al (2003) propose a two-stage model for the acquisition of language-specific phonetic categories in infants by the age of 6 – 8 months. They suggest that ‘auditory learning’ relies on an innate ability to distribute sound stimuli along a scale of categories. Acoustic information that is heard by an infant will be weighted by the brain towards one category or another. These categories are not absolutely discrete, but the sounds of each of the commonly heard vowel sounds will be perceived to cluster around a particular place on the scale. The sound that is heard by the infant, or modelled by the neural net, will be perceived as being closer to one or the other area and categorised accordingly. Eventually this will lead to discrete categories being formed that will then allow the infant to move on to learn lexical distinctions. The important implication of this theory for the present research is that Boersma et al. suggest that more common sounds, such as vowel sounds, provide a perceptual ‘magnet’ to help the infant distinguish specific patterns that are most common in the language that they hear people speak every day.

3.1.6: Vowel use in nursery rhymes and songs

There are many regular features of nursery songs that are relevant when looking at how infants might develop their communicative abilities, and rhyming vowel sounds appear to be one of the most salient. Milligan et al (2003) state that lullabies and play songs can be considered to be “musical analogues of motherese” (p. 3). Patel and Pertetz (1997) state that “song is a universal form of auditory expression in which music and speech are intrinsically related” (p. 206), and they suggest that song and speech may be represented in the brain in a similar manner. However, they do point out that there are some interesting differences in some features of speech and song. For example, there are differences in the ways in which pitch scales are utilised in music and in speech. But there are correspondences in the ways that both musical and linguistic stimuli are processed by the brain, and Patel and Pertetz (1997) suggest that although music and language are not interdependent, they may share some underlying functional mechanisms.

Similarities between features of speech and song may be more pronounced for vowel sounds because they are more commonly used to vary pitch. Vowels in nursery songs

are used in the same way as in infant-directed speech; they are elongated for the exaggerated pitch shifts that occur during infant directed singing (Trainor et al, 2000). Infants show a preference for infant-directed singing of nursery songs over adult-directed versions of the same songs, from birth (Ikeda and Masataka, 1999; Masataka, 1999; Trehub and Nakata, 2002). Timing, structure and rhythm appear to be utilised in a very similar manner to attract the interest of infants in many different cultures. Nursery songs and rhythmic vocal games can reduce infant stress (Trehub and Nakata, 2002). They are similar to infant directed speech in that they have a simple, repetitive structure. Their intonation is exaggerated in a manner that helps define the boundaries of rhythmic units that are equivalent to syllables and phrases. They create a predictable sound medium, whereby infants can begin to anticipate and imitate acoustic features. The features of nursery songs and rhythmic vocal games appear to be adapted to facilitate the passing on of cultural messages, and it may be that they also provide parents with a means by which they can voice their fears and worries for their own satisfaction, while sharing the music with their infants.

Nursery songs and rhythmic vocal games clearly can create a link between pre-verbal communication and the development of language and it may be that vowel sounds play a central role in creating this link. They provide vocalisations with a varied playfulness. This could have a functional purpose for the infant in that they are given an attractive and memorable structure by imitation of which they can practise and expand their own vocal range. More than this, however, these songs and games allow humans to be involved in activities that are fun and playful. Music and play therapists know of the therapeutic benefits of this type of enjoyable activity (Jernberg and Booth, 2001). Playful interactions would appear to increase the opportunities for mutually satisfactory interactions because they require a close coordination of timing, purpose and understanding between partners (Reddy, 2002).

Rhythmic musical expression in motherese and in the emotional narratives of nursery games has been shown to have dual function in communication with infants. Mothers use nursery rhymes to regulate 'state of arousal' or 'alertness', and 'stress', promoting a mutually satisfying relationship between infant and adult that aids learning and

brain development (Shimura and Imaizumi, 1995; Trehub et al, 1997; Rochat, 1999). Infants have been shown to have episodic memory for musical stimuli that relates both to individual acoustic aspects and more general structural characteristics such as phrase endings (Trehub et al, 1997; Trehub and Nakata, 2002; Trainor, Palmer et al, 2001; Mattys and Jusczyk, 2001).

In subsequent months, toward the end of the first year and beyond, nursery songs and rhythmical games attract the infant's awareness and thinking to joint cognitive interest in actions and objects, facilitating the learning of culture-specific rituals and skills, including language (Stern and Gibbon, 1980; Trevarthen and Aitken, 1994; Morgan and Saffran, 1995; Trehub et al, 1997; Kuhl, 1998; Jaffe et al, 2001; Rochat, 1999).

There is evidence that infants are attracted first to the regular timing, repetition and affective quality of maternal speech, and that they soon learn to recognise certain often repeated expressive forms in rituals of play and in speech, including ones that identify affective state, parental culture and language (Papoušek, 1996; Trevarthen and Aitken, 1994; Lecanuet et al, 1996; Trehub et al, 1997; Friederici and Wessels, 1993; Mehler, 2000).

So, the literature reviewed here suggests that acoustic features of voice can provide a medium in which infants can segment speech, can discriminate their native language, can identify speaker identity and gender and can understand something about the affective states of speakers. However, the literature does not address the specific issue of how physical acoustic properties of speech may provide a means by which emotional meanings can be shared between a parent and their pre-linguistic infant. This review highlights several acoustic components that are worthy of further study (including pitch, intensity and rhythm/tempo) by means of which emotion can be mutually understood in parent-infant communication. The review provides a framework for examining those acoustic features of vowel sounds that have been identified as particularly salient for infants.

3.1.7: Research Questions

This study was carried out to explore the following research questions:

1. What role do vowel sounds play in the communication of a mother's affect?
2. How do infants react to affective changes in the mother's voice?

An 'emotional voice' paradigm was developed and implemented to investigate if there are systematic relationships between acoustic features of mothers' vowel sounds and the specific emotional intent of their vocalisations. The paradigm explored whether or not the infant noticed changes in the quality of the mother's voice when she was required to change affective features of her speech. The assumption of the model is that if the mothers altered the pitch, intensity and duration of their vowels to convey specific emotions, and if the infants showed a change in behaviour in relation to this, prosodic modulation of vowels is a principal component for perception or expression of the emotional narrative by infants. The experiment was carried out to provide a description of how infants behaved in response to affective changes in the mother's voice and to see whether the infants' affective expressions corresponded to emotions expressed in the mother's voice.

It was expected that acoustic features of vowel sounds in mothers' voices would be different when the mother spoke to her infant in different emotional voices ('happy', 'sad' and 'bored'). It was hypothesised that infants would be sensitive to changes in emotional expression, and that they would show this sensitivity through matching their own expressive gestures to those of their mother in each emotional condition. It was predicted that infants would show distress, confusion, surprise or amusement in reaction to changes in the emotions expressed by their mother. If infants respond in this way to changes in the mothers' voices, this would directly test the theory that infants are sensitive to emotional aspects of acoustic features of mothers' speech. This would provide evidence to support the idea that acoustic information and in particular the information contained within vowel sounds is necessary for affective exchange during mutually satisfying communication, at least for hearing infants.

3.2: Method

3.2.1: Participants

Eleven mothers and their infants took part in the Study. There were six female infants aged between 5 months and 9 months (mean age 7.2 months) and five male infants aged between 3 months and 8 months (mean age 5.4 months). The mean age of the infants was 6.7 months.

3.2.2: Procedure

Mothers and infants were invited to visit the laboratory. Each infant was placed in a child seat facing the mother. The purpose of the study was explained to the mothers, and each was told she would be asked to speak, not sing, a familiar nursery rhyme, *Round and Round the Garden*, to her baby.

Each mother was asked to say the rhyme in the following four emotional conditions, and in this order:

Condition 1 = 'Happy' voice (1)

Condition 2 = 'Sad' voice

Condition 3 = 'Happy' voice (2)

Condition 4 = 'Bored' voice

The second 'Happy' Condition was included to ensure that the infant had the chance to recover from any emotional reaction to the 'Sad' Condition. The infant's reactions were video taped. The mother was not filmed but a recording of her voice was taken from the videotape. Two types of analysis were carried out:

- Acoustic analysis of the mothers' vocal expressions,
- Rating of non-vocal responses of the infant to their mother's voice.

An essential part of the procedure was that the mothers had to say the rhyme several times, each time as if they were expressing a different emotion. They had to 'act'

each required emotional condition, so it was necessary to examine how this influenced the infants' responses and so, indirectly, the raters' coding of emotional matching between mother and infant.

When mothers had finished saying the rhyme to their child, they were asked to rate each emotional condition for how difficult it was to 'act'. Mothers rated each emotional condition on a six point scale, where a score of 1 was very difficult to act and a score of 6 was very easy to act. Their ratings for each condition were compared.

3.2.3: Acoustic Analysis of the Mothers' Vocal Expressions

Acoustic analysis was carried out on the mothers' voices using Praat phonetic analysis software, to see if there were commonalities in acoustic characteristics of the mothers' voice in each emotional condition. A general acoustic analysis was carried out on all vocalisations to provide a description of pitch and intensity information for each emotional condition. Vowel sounds were then isolated and classified using visual spectrographic output (showing vowel nuclei and formants) calculated by the Praat programme. An acoustic check was carried out to ensure that only the vowel sounds had been isolated.

Speech sounds (such as semivowels) and transitional sounds leading into vowels (glides) that have a similar quality to vowel sounds but that can also operate as consonants were treated in the following way: those preceding the vowel (pre-vocalic) sound such as 'qu' (as in the word 'queen') were treated as consonants, whereas those coming after the vowel sound (post-vocalic) such as 'ow' (as in the word 'how') were classified as vowels.

Vowel length was measured and vowels were then categorised as being 'Extended' vowels (if they were over 250ms), 'Long' vowels (if they were between 151-250 ms) and 'Short' vowels (if they were between 50-150 ms). All vowel sounds that measured less than 50ms were disregarded from the analysis because these vowels

were so brief that they were often indistinguishable or could not be isolated from the consonants surrounding them.

Pitch range, intensity and duration were then calculated by PRAAT for each vowel sound produced by the mother. The Praat programme could not calculate pitch information for whispered speech and so although whispered vowel sounds were categorised and measured for intensity and duration, they were not included in the pitch analysis. The vowel analysis explored whether there were shared features in how pitch, intensity and duration of vowel sounds were utilised by mothers to express the specific emotions in each condition.

3.2.4: Rating of Non-Vocal Responses of the Infant to their Mother's Voice

The video recordings of infant reactions in each condition were arranged in a randomised sequence. Thirty-three recordings – one for each of the eleven infants, in each of the 'Happy 1', the 'Sad' and the 'Bored' conditions were played to three raters (two female non-parents and one male parent) who were blind to the purpose of the test. Recordings of the nursery rhyme varied in length between mothers, and Table 3.1 provides overall information about mean length of the rhyme in different emotional conditions. On average the mothers took longer to say the 'Sad' version of the rhyme.

Table 3.1: Mean length of rhyme in each emotional condition

	<i>Happy</i>	<i>Sad</i>	<i>Bored</i>
Mean length of rhyme (in seconds)	7.782	9.819	8.017

The raters watched the recordings in a group, but were asked not to discuss their ratings. They did not hear the sound from the videos and they were given the following instructions:

“You will see some video clips of infants listening to their mothers. You can ask to see the clip again if you need to. Please make a judgement about how you think the infant in each clip is reacting, choosing an option from the following list: (1) the baby

is 'happy', (2) the baby is 'sad', (3) the baby is 'bored'. In the box, please describe any of the infant's gestures, behaviours or expressions that led you to describe the infant reaction in this way".

Picture 3.1: One infant's reaction to the 'Sad' condition



Raters' judgments were counted as 'Matching' when the judgement about emotion corresponded to the emotion that the mother was expressing (an example is provided in Picture 3.1). 'Matching' judgements were summed for each emotional condition. The words that each rater used to describe the infants' non-vocal responses in each emotional condition were compared.

3.3: Results

It was predicted that there would be features of the mother's vowel sounds that would be characteristic for specific emotional conditions. It was expected that infants would be sensitive to manipulations of emotional expression in their mothers' voices and that they would show this sensitivity through differences in their expressive gestures between each of three emotional conditions ('Happy', 'Sad', and 'Bored'). To test these ideas, the analysis of Study 2 was organised to collect data on the following:

- The emotional characteristics of the mothers' vocal expression in each emotional condition,

- Ratings of infant non-vocal reactions to the mother's voice,
- The levels of difficulty experienced by the mothers in 'acting' different emotions.

3.3.1: Dimensions of Emotion in the Mother's Voice

3.3.1.1: Pitch range of all vocalisations

To understand whether or not there were characteristic features of the mother's voice that could convey specific emotions to the infant, pitch analysis of the mother's voice in each condition was carried out using the Praat acoustic analysis computer programme. The program was used to isolate and measure the minimum and maximum pitch of the mother's voice during each emotional condition, providing a pitch range for each mother. This analysis first combined all vocalisations to explore whether there were differences in the mean pitch between conditions, and further analysis was then carried out on isolated vowel sounds. If any difference was established for the isolated vowel sounds, it would be assumed that vowel sounds present key acoustic features serving the vocal expression of affective state.

Table 3.2: Pitch range of all vocalisations, for each mother in each emotional condition (measured in Hertz)

Infant I.D*	<i>Mean of Pitch Range (Hz) in Each Condition</i>			
	'Happy 1'	'Sad'	'Happy 2'	'Bored'
1 (F.1)	395.75	207.19	303.28	232.56
2 (F.1)	399.97	377.30	403.59	255.73
3 (M.1)	399.76	205.34	338.44	129.16
4 (M.1)	408.45	366.10	396.33	397.88
5 (M.1)	314.51	222.47	314.78	271.94
6 (M.1)	297.44	409.75	381.86	271.94
7 (F.2)	348.46	142.25	408.98	214.70
8 (F.2)	300.56	364.26	306.43	361.76
9 (F.2)	396.81	315.14	305.01	370.90
10 (F.2)	395.56	300.07	380.73	194.06
11 (M.2)	398.89	245.23	387.61	159.65
Mean pitch	368.74	286.91	357.01	260.03
range				
Standard	44.44	87.44	43.36	87.36
Deviation				

Note*

The letter identifies the gender of each infant (F=female, M=male),

The number after the decimal point shows age (1 = under 6 months , 2 = over 6 months)

Table 3.2 shows that there is little difference in the means of the two 'Happy' conditions. The recitations in both 'Happy' conditions have a similar pitch, which is much higher than those of either the 'Sad' or the 'Bored' conditions. From this point on, pitch analyses excluded the 'Happy 2' condition. The standard deviations suggest that although there is some individual variation, between mothers, there are clear differences in the 'Sad' and 'Bored' conditions, suggesting that the main differences may lie between positive and negative emotions. Graph 3.1 shows the distribution for the pitch range of all vocalisations in each condition. The graph shows that mean pitch ranges for the different conditions are sufficiently different to merit further analysis.

Graph 3.1: Distribution for pitch range of all vocalisations in each condition



A repeated-measures ANOVA with emotional condition as a within-subjects factor revealed that pitch range differed significantly between the emotional conditions ($F(2, 20) = 7.426, p=0.04$). This finding suggests that pitch range in vocalisations may be an important measure of the expression of different emotions. A paired samples t-test revealed that there was a significant difference between the 'Happy' and 'Sad' conditions ($t=2.600, df=10, p=.027$), a significant difference between the 'Happy' and 'Bored' conditions ($t=3.386, df=10, p=.007$) but no significant difference between the

‘Sad’ and ‘Bored’ conditions ($p=.286$). This suggests that the difference in pitch range lies mainly between positive and negative emotional conditions.

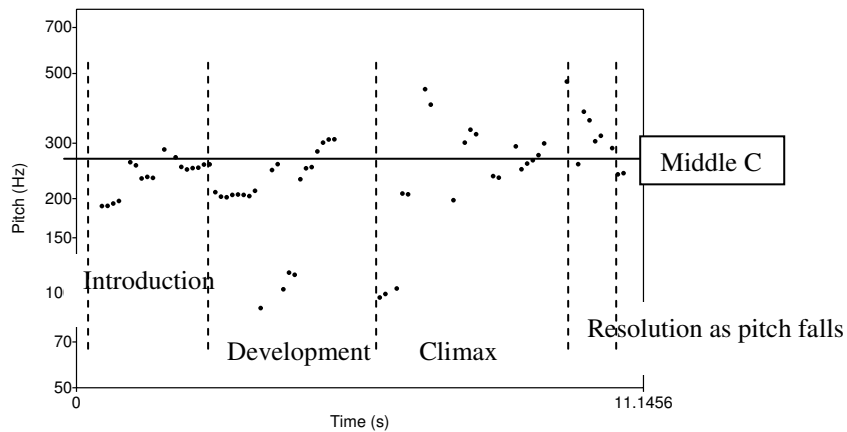
Pitch plots were produced using the Praat programme and they provide a useful overview of how pitch changes in each emotional condition. Graphs 3.2 to 3.4 (next page) illustrate a typical example of this for one mother. In these graphs pitch is shown (on the vertical axis) as a function of time and is represented by a logarithmic scale. In the ‘Happy’ condition (Graph 3.2, next page) the mother’s pitch ranges from 76.427 Hz to 471.178 Hz. Stephen Malloch (1997) states that it is important to remember that human perception of sound is translated in the brain into a psychoacoustic experience, which has more in common with musical scales as opposed to physical measures of frequency. Trevarthen (2007 personal communication) proposes that Middle C (261.63Hz) provides “a useful horizon for distinguishing positive and negative emotions.

The narrative pattern of her voice exhibits an introduction phase where she sets up the rhythm for the rhyme, a development phase where she builds the pitch and lets it fall and build again and a climax phase where her pitch increases to its maximum level, before the resolution phase where her pitch falls at the end of the phrase. This pattern is commonly found in nursery rhymes and appears to be used by the mothers to maintain the infant’s interest and to create excitement.

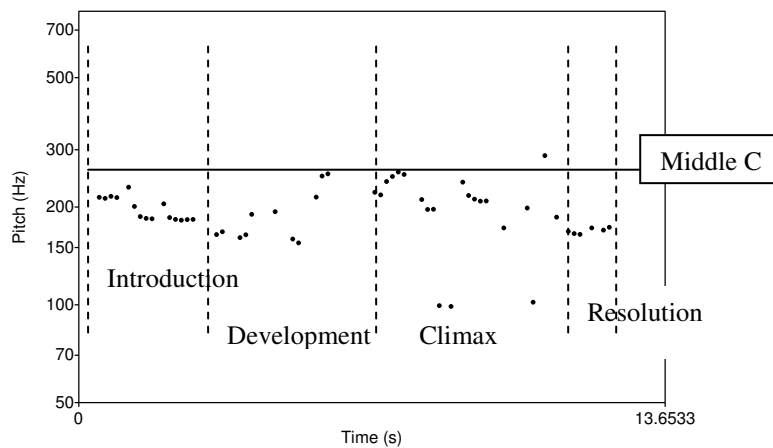
In the ‘Sad’ Condition (Graph 3.3, next page) it can be seen that the mother’s pitch is much lower and has a much smaller variation (81.309 Hz to 287.499 Hz). The intonation of the mother’s voice is much flatter, showing less expression and the beginning phrase does not build in the same way. In this condition, the mother slows down her speech and the rhyme takes longer to say and there is no falling pitch at the end of the phrase and her voice falls below Middle C. In the ‘Bored’ Condition (Graph 3.4, next page) the mother’s voice again remains in a much lower pitch, ranging from 75.907 Hz to 297.710 Hz (a slightly greater range than in the ‘Sad’ Condition but still below the Middle C). The beginning phrase of the rhyme falls steadily and the second part of the rhyme is said in a ‘sighing’ voice, which accounts

for the way in which the pitch plot appears to break up towards the end of the rhyme. The mother vocalises below Middle C. In this case a falling pitch is evident at the end of the phrase but is incorporated into the singing voice.

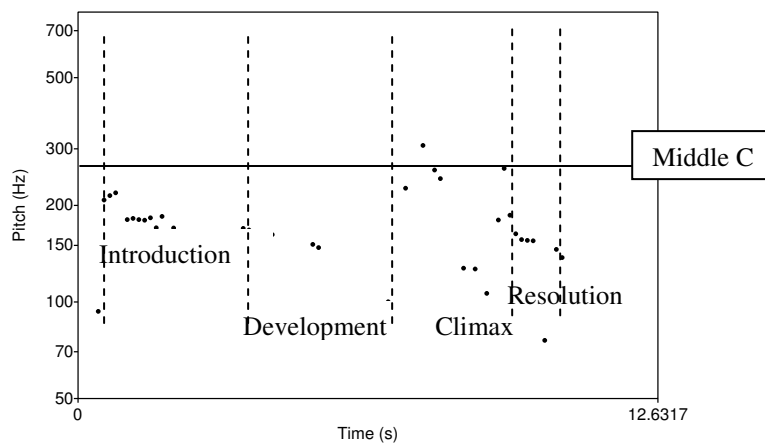
Graph 3.2: Pitch of one mother's voice in the 'Happy' Condition



Graph 3.3: Pitch of one mother's voice in the 'Sad' Condition



Graph 3.4: Pitch of one Mother's voice in the 'Bored' Condition



3.3.1.2: Intensity range of all vocalisations

Praat acoustic analysis software was also used to analyse the intensity range of the mothers' voices (measured in decibels). Again this was a general analysis on both vowels and consonants. Precise measurements of minimum and maximum intensity levels were taken for the mother's voice during each emotional condition. Table 3.3 provides descriptive information about the intensity range for each mother.

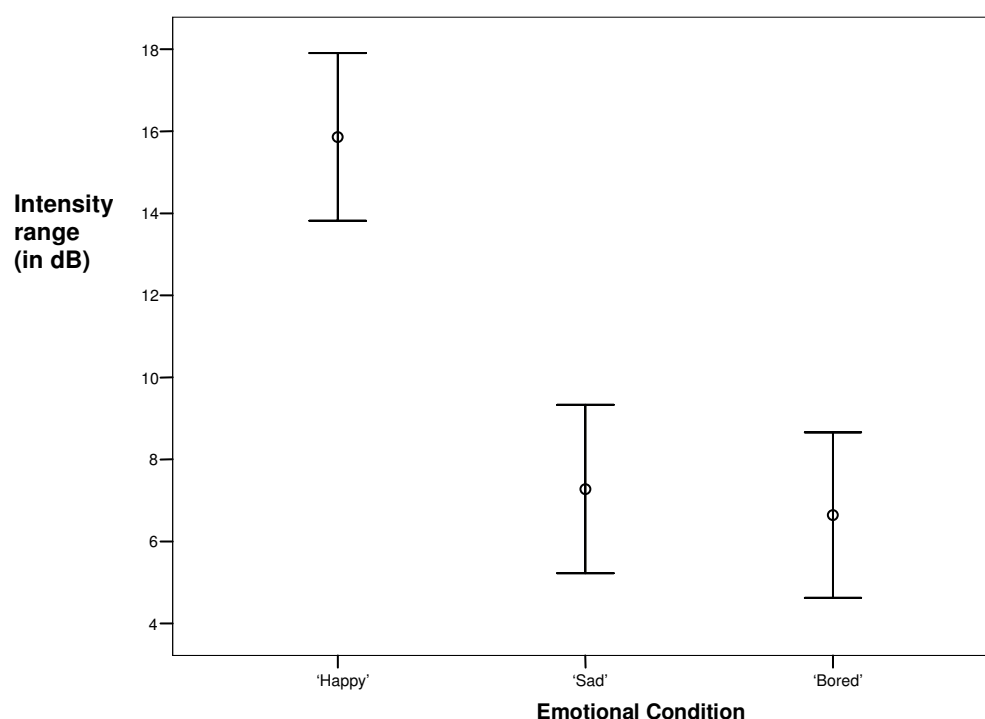
Table 3.3: Intensity range for all vocalisations for each mother in each emotional condition (measured in Decibels)

Infant I.D	<i>Intensity range (dB) for each Condition</i>			
	'Happy 1'	'Sad'	'Happy 2'	'Bored'
1 (F.1)	18.76	10.05	19.91	11.17
2 (F.1)	15.39	7.57	20.55	7.47
3 (M.1)	15.99	4.33	6.15	3.29
4 (M.1)	20.56	9.52	14.10	11.76
5 (M.1)	10.32	7.47	15.03	5.14
6 (M.1)	17.54	6.55	17.83	6.78
7 (F.2)	17.84	4.01	13.84	8.01
8 (F.2)	14.63	13.87	10.81	2.83
9 (F.2)	11.22	6.08	13.54	6.28
10 (F.2)	15.23	7.30	13.45	7.21
11 (M.2)	16.98	3.30	13.81	3.14
Mean Intensity	15.86	7.28	14.46	6.64
range				
Standard	3.05	3.06	4.05	3.00
Deviation				

The table shows that as with the pitch, the intensity is similar for both of the 'Happy' conditions and so from this point on, intensity analysis will be carried out using data from the 'Happy 1' condition only. The 'Happy' conditions have a much wider intensity range than either the 'Sad' or the 'Bored' conditions. The standard deviations suggest that this appears to be fairly consistent across the mothers and inspection of the individual data confirms this.

The distribution of voice intensity range is shown in Graph 3.5 and indicates that the distribution is sufficiently different to merit further analysis. The distribution for the 'Happy' condition is much higher than for the 'Sad' and 'Bored' conditions which appear to be very similar.

Graph 3.5: Distribution for intensity range of all vocalisations in each condition



A repeated measures ANOVA was carried out with emotional condition as a within-subjects factor and this confirmed that there was a highly significant effect of emotional condition on intensity range, ($F(2, 20) = 42.500, p = .0001$).

To explore whether there was a significant difference between each of the emotional conditions, paired samples t tests were carried out and showed that there was a significant difference between the 'Happy' and 'Sad' conditions ($t = 6.678, df = 10, p = .0001$) and between the 'Happy' and 'Bored' conditions ($t = 10.578, df = 10, p = .0001$). However, there was no significant difference between the 'Sad' and 'Bored' conditions ($t = .549, df = 10, p = .595$ ns). Again it seems that intensity range is utilised to distinguish between positive and negative emotions.

Using Praat acoustic software involves a visual examination of spectrographic information. During these examinations it became apparent that a great deal of the variation of the pitch range and intensity was occurring in the vowel sounds of the mothers' speech. To understand more fully if this was the case, a more detailed

examination of all vowel sounds was carried out to describe the following features of these sounds:

- Pitch range,
- Absolute intensity,
- Vowel length (duration).

3.3.1.3: Pitch Range of Vowels

Vowel sounds were identified and isolated using Praat acoustic software. Visual inspection of spectrographs (including formant information) allowed vowel nuclei to be precisely measured. Maximum and minimum pitch measurements and pitch range were calculated for all of the isolated vowel sounds for each mother in each condition. Means were calculated for all pitch information for each mother and Table 3.4 provides detailed description of the results of the pitch analysis.

Table 3.4: Mean pitch range (in Hz) for vowel sounds for each mother in each condition

Infant ID	<i>Mean pitch range (in Hz)</i>		
	‘Happy’	‘Sad’	‘Bored’
1 (F.1)	45.76	20.97	14.67
2 (F.1)	69.77	38.26	6.91
3 (M.1)	32.93	20.52	23.99
4 (M.1)	60.07	54.34	24.63
5 (M.1)	30.94	52.58	52.83
6 (M.1)	88.56	52.40	19.18
7 (F.2)	33.94	14.88	17.91
8 (F.2)	53.03	61.10	9.66
9 (F.2)	57.51	12.69	16.32
10 (F.2)	34.67	34.52	20.90
11 (M.2)	61.18	16.41	16.71
Mean pitch range for condition	51.67	34.42	20.34
SD for condition	18.18	18.25	12.05

The table shows that there are differences in the mean pitch range for vowels in each condition. Although there are differences in the mean pitch range of individual mothers, the mean pitch range for conditions shows that the ‘Happy’ condition has the highest range, followed by the ‘Sad’ condition and lastly the ‘Bored’ condition. Standard deviations for each condition show that variation in pitch range is similar for the ‘Happy’ and ‘Sad’ conditions but much lower in the ‘Bored’ condition.

A repeated-measure ANOVA showed that there were indeed significant differences between conditions with respect to the mean pitch range for vowels, ($F(2,20)=10.705$, $p=.001$). To explore the differences in more detail, paired samples t-tests were carried out and showed that there was a significant difference between the ‘Happy’ and ‘Sad’ conditions ($t=2.627$, $df=10$, $p=.025$), between the ‘Happy’ and ‘Bored’ conditions ($t=4.000$, $df=10$, $p=.003$) and between the ‘Sad’ and ‘Bored’ conditions ($t=2.432$, $df=10$, $p=.035$). So it seems that the pitch range of vowel sounds is an important acoustic (physical) feature which allows expressive and perceptual differentiation of emotional (psychological) states.

3.3.1.4: Intensity of vowel sounds

The intensity of all vowel sounds was measured to see if mothers used differences in intensity to express specific emotions. Table 3.5 shows the results of this measurement:

Table 3.5: Intensity measurements in Decibels (dB), for all vowels in each emotional condition

Infant ID	<i>‘Happy’</i>		<i>‘Sad’</i>		<i>‘Bored’</i>	
	Mean intensity	SD intensity	Mean intensity	SD intensity	Mean intensity	SD intensity
1 (F.1)	67.12	3.43	63.08	1.50	63.33	1.74
2 (F.1)	67.67	5.95	57.96	3.92	53.03	5.33
3 (M.1)	57.88	8.39	50.09	3.56	53.28	4.06
4 (M.1)	70.09	4.80	64.37	4.18	64.67	4.64
5 (M.1)	62.36	3.97	60.40	4.48	57.79	4.07
6 (M.1)	70.23	3.64	63.00	4.26	56.60	4.33
7 (F.2)	63.21	1.17	62.23	0.71	53.62	7.46
8 (F.2)	65.03	4.73	55.62	3.10	48.92	2.96
9 (F.2)	62.67	5.50	54.18	4.60	54.58	4.66
10 (F.2)	60.17	6.00	54.21	5.49	54.40	4.04
11 (M.2)	64.10	5.22	52.22	3.78	53.75	2.87

	‘Happy’	‘Sad’	‘Bored’
Mean intensity for condition	64.59	57.94	55.81
SD for condition	3.92	4.96	4.69

The table shows that again there do seem to be some differences in the overall mean intensity for each emotional condition. The standard deviations reveal that there is similar variance in each condition and there seems to be a general pattern where intensity is louder for the ‘Happy’ condition and is most subdued in the ‘Bored’ condition. However, again there are some individual variations in the patterns of intensity that mothers use to express the negative conditions.

A repeated measures ANOVA with emotional condition as the within-subjects factor showed that there was a significant difference between emotional conditions ($F(2, 20)=30.035, p=.0001$). Paired samples t-tests revealed that there was a significant difference in intensity between the ‘Happy’ and ‘Sad’ conditions ($t=6.605, df=10, p=.0001$) and between the ‘Happy’ and ‘Bored’ conditions ($t=6.588, df=10, p=.0001$) but not between the ‘Sad’ and ‘Bored’ conditions ($t=1.798, df=10, p=.102$ ns). It seems that mothers lower the intensity of their vowel sounds when expressing negative emotions.

3.3.1.5: Duration of mothers’ vowels in each condition

All of the mothers’ vowels were measured for duration and categorised according to the following definitions:

‘Short’ vowels: 50ms-150ms

‘Long’ vowels: 151ms-250ms

‘Extended’ vowels: over 250ms

Table 3.6 shows the frequency of each vowel type in each emotional condition for all mothers. It shows that the ‘Sad’ condition has more ‘Extended’ vowels than the other

conditions, that there are more ‘Long’ vowels in the ‘Sad’ condition and more ‘Short’ vowels in the ‘Happy’ and ‘Bored’ conditions than in the ‘Sad’ condition.

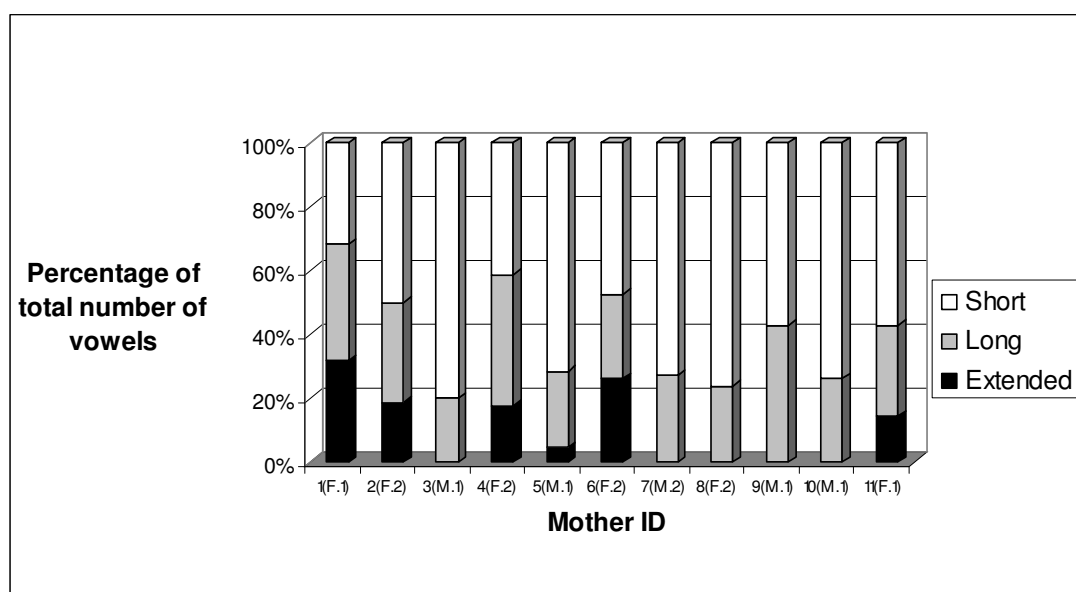
Table 3.6: Frequency of vowel type in each emotional condition

	<i>Frequency of vowel types in each condition</i>		
	Happy	Sad	Bored
Extended	22 (27.2%)	35 (43.2%)	24 (29.6%)
Long	65 (32.3%)	88 (43.6%)	49 (24.3%)
Short	132 (35.1%)	103 (27.4%)	141 (37.5%)
Total	219 (33.2%)	226 (34.3%)	214 (32.5%)

Percentages show that there are more ‘Extended’ and ‘Long’ vowels in the ‘Sad’ condition than in the ‘Happy’ or ‘Bored’ conditions, more ‘Short’ vowels in the ‘Happy’ and ‘Bored’ conditions . All vowels were entered into a 3 (type) x 3 (emotion) chi square analysis, which showed that there was a significant overall difference in the frequency of vowel types between emotional conditions ($X^2=20.85$, $df=4$, $p=.0001$).

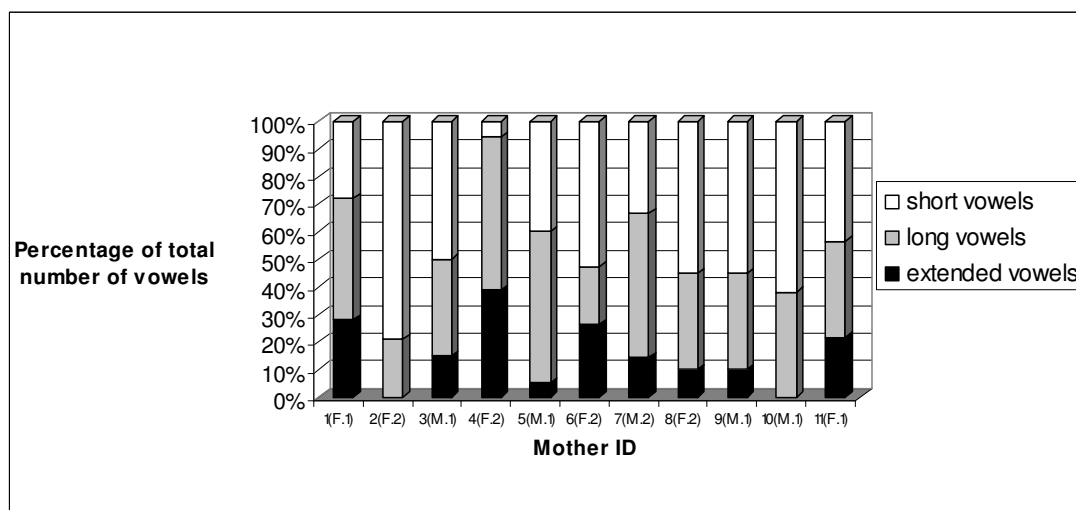
Graphs 3.9, 3.10 (see next page) and 3.11 (see page 109) show the vowel types for each mother in each condition. Because the mothers are speaking the same rhyme with the different emotions there is a natural constraint on how many vowels will be used. However, the figures show that there is individual variation in the ways that mothers extend and shorten their vowels to express each emotion while they speak. This variation suggests that vowels are used differently by the mothers in their individual speech styles to express the different emotions.

Graph 3.9: Vowel Types for each mother in the ‘Happy’ Condition



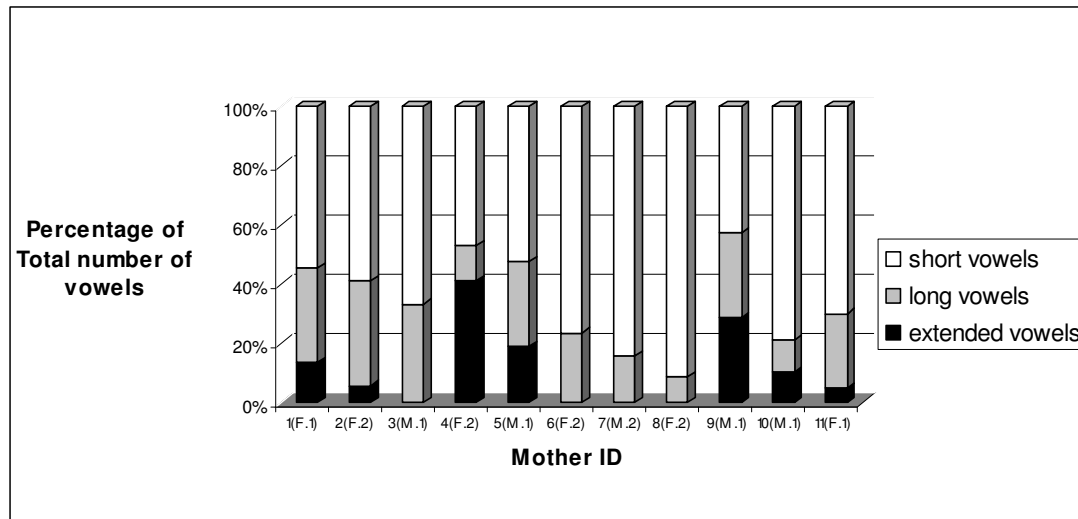
In the ‘Happy’ condition, five out of eleven mothers did not use any ‘Extended’ vowels, but two others used ‘Extended’ vowels in over 20% of their vocalisations. Some of the mothers used ‘Extended’ vowels as a means of creating surprise and anticipation, or as a way of introducing humour. Those mothers who did not use ‘Extended’ vowels at all often appeared to speak more quickly and used more ‘Short’ vowels to give their rhyme a bouncy feel. Again this seemed to be used to maintain the infants’ interest.

Graph 3.10: Vowel Types for each mother in the ‘Sad’ Condition



All but two of the mothers used extended vowels in the ‘Sad’ condition. This appeared to be a common way for the individual mothers to express sadness in their voice. Some of the mothers (for example 5M.1) used a much higher percentage of ‘Long’ vowels in the ‘Sad’ condition in comparison to the ‘Happy’ condition. In general mothers in the ‘Sad’ condition seemed to use ‘Extended’ vowels of low pitch and intensity to give the rhyme a mournful sound.

Graph 3.11: Vowel Types for each mother in the ‘Bored’ Condition



In the ‘Bored’ condition 7 out of 11 of the mothers used ‘Extended’ vowels and for some mothers the percentage of ‘Long’ vowels seems to be less than for the ‘Sad’ condition. Again, ‘Extended’ vocalisations had a depressed tone rather than the playful one used in ‘Extended’ vowels by some mothers in the ‘Happy’ performance of the verse.

3.3.1.6: Whispered vowels and infants’ vocalisations

Mothers also used a whispered voice to express specific emotions, although the overall frequency of these whispered vowels was low. Whispered vowel sounds were measured for duration but it was not possible to gain pitch information for them as their frequency was too high for pitch to be measured consistently. Table 3.7 (next page) shows the frequency of whispered vowels in each condition.

Table 3.7: Whispered vowels in each condition

	<i>Happy</i>	<i>Sad</i>	<i>Bored</i>
Total frequencies for whispered vowels	1	10	12

Clearly there are more whispered vowel sounds in the ‘Sad’ and ‘Bored’ conditions and indeed it seems that whispered vowel sounds were used by some mothers in combination with a sighing, or broken voice to express sadness and boredom.

Another feature that was noticeable from the vowel analysis was that infants occasionally vocalised concurrently with the mother when she was saying the rhyme. Table 3.8 shows that infants sometimes vocalised at the same time as their mother in the ‘Sad’ and ‘Bored’ conditions, whereas they did not do this at all in the ‘Happy’ condition.

Table 3.8: Concurrent infant vocalisations in each emotional condition

	<i>Happy</i>	<i>Sad</i>	<i>Bored</i>
Total frequencies for concurrent infant vocalisations	0	5	4

3.3.2: Infant reactions to mother’s voice

The second area chosen for analysis focused on the following question: Do infants’ responses to their mothers’ voices vary according to emotional condition?

If raters judge that infants’ gestures and expressions ‘match’ or ‘mirror’ the emotion in their mother’s voice then this would suggest the following:

Infants respond to specific emotions in a sympathetic or empathetic way. They mirror the mother’s affective state with a matching emotional gesture, suggesting at the very least that they are sensitive to subtle emotional changes in the mother’s behaviour.

Infants respond to their mother's affective state in ways that can be clearly understood by raters who are blind in that they do not have access to any other contextual information, other than the infants' gestural and expressive behaviour.

Three raters (who were blind to the purpose of the study) coded each of the eleven infants' reactions to their mothers' voices in four emotional conditions; 'Happy 1', 'Sad', 'Happy 2' and 'Bored'. Each time the raters' coding of infant gestures corresponded with the emotional situation, it was counted as a 'Matching' score. 'Matching' scores for the 'Happy 1' and 'Happy 2' conditions were identical and so from this point on, analysis is carried out using only the data from the 'Happy 1' condition. 'Matching' scores were calculated out of a maximum possible of 33 (3 raters x 11 infants, in each condition). Table 3.9 shows the number of raters' codings that 'Matched' infant gestures with emotional situations.

Table 3.9: 'Matches' for each condition

	<i>Frequency (percentage) of judgement for each condition</i>	
	'Matching'	'Non-matching'
Happy	18 (54.4%)	15 (45.5%)
Sad	4 (12.1%)	29 (87.9%)
Bored	20 (60.6%)	13 (39.4%)
Overall	42	57

The table shows that overall participants made more 'Non-matching' judgements than they did 'Matching'. However there did seem to be differences within emotional conditions. 54.5% of raters' judgements were 'Matching' in the 'Happy' condition, and 60.6% in the 'Bored' condition. The lowest number of 'Matches' was for the 'Sad' condition (12.1%) suggesting that infant responses to this condition were much less likely to mirror the emotion expressed by the mother. A 3x2 ('Happy'/'Sad'/'Bored' x 'Matching'/'Non matching') Chi-Square analysis showed that there was a significant association between emotional condition and the frequency of 'Matching' responses ($X^2=18.857$, $df=2$, $p=.0001$). Separate goodness-of-fit Chi-squares were carried out for each emotional condition using expected frequencies of 16.5 (which would assume the Null hypothesis because 'Matching' and 'Non-matching judgements equalled 33 in each emotional condition). This analysis showed that there was no significant difference between 'Matching' and 'Non matching' judgements in the 'Happy' condition ($X^2=.273$, $df=1$, $p=.602$ ns) and the

‘Bored’ condition ($X^2= 1.485$, $df=1$, $p=.223$ ns) but that there were significantly less ‘Matching’ than ‘Non matching’ judgements in the ‘Sad’ condition ($X^2= 18.939$, $df=1$, $p=.0001$).

Table 3.10 below shows that when all the raters’ codings were counted, there were far fewer ‘Sad’ descriptions of infant gestures.

Table 3.10: Total number of codings given by each rater for each emotion (including both ‘Matches’ and ‘Non-matches’)

<i>Rater</i>	<i>‘Happy’</i>	<i>‘Sad’</i>	<i>‘Bored’</i>
1	17	5	13
2	15	2	18
3	19	3	13
Totals	51	10	44

It is not possible, from the results of Study 2, to say why raters were less likely to code infant behaviour as being ‘Sad’. The fact that all the raters have a similar pattern of coding suggests that it is the infants’ behaviour that is affecting coding, as opposed to reluctance on the part of the raters to code infant behaviour as ‘Sad’.

Another area for analysis was to explore whether some infants are easier to ‘read’ than others. There could be many reasons for this. Some infants may be more expressive than others, and may be more likely to match or mirror their mother’s emotional state. To see if individual differences in infant gestures and expressions had any effect on the findings, ‘Matches’ for each infant were counted. Table 3.11 (next page) provides details of this count. The highest possible total for each infant was 9 (3 raters coding infant gestures to 3 emotional versions of the rhyme).

Table 3.11: Total ‘Matches’ (out of 9) for each infant

<i>Infant</i>	<i>Infant Age (months)</i>	<i>Total of ‘Matching’ judgements (out of 9)</i>
1(F.1)	5	5
2(F.1)	6	3
3(M.1)	6	8
4(M.1)	4	3
5(M.1)	3	0
6(M.1)	6	4
7(F.2)	8	5
8(F.2)	7	2
9(F.2)	9	4
10(F.2)	9	5
11(M.2)	7	3

The table shows that there appear to be differences between the matching coding for individual infants. For example, one infant had no ‘Matches’ whereas another had 8. Coders, indeed, commented that some infants’ expressions were much more difficult to ‘read’. Infants whose gestures and expressions were difficult to read tended to be acting more interested in their surroundings, possibly avoiding their mothers, and this made it more difficult for the raters to decide how the infant was reacting to their mother saying the rhyme. Some infants were more difficult to read because they looked either very happy or very sad throughout all of the conditions, and their gestures and expressions did not seem to alter much during the test. A chi square showed that there was a significant association between the individual infant being coded and raters’ ability to make a match ($X^2=30.417$, $df=10$, $p=.001$). So it seems that individual differences are evident, although it is not clear from this study whether these differences are in infants’ expressive behaviour, or in their responsiveness to their mother.

3.3.3: Level of difficulty in ‘acting’ emotions

The third area for analysis was to examine whether mothers found some emotions harder to ‘act’ than others and to see whether this was associated with the raters’ coding of emotional matching. It was important to take account of this for two related reasons:

It may be that the infants can sense that the mothers' emotional expression is not real, and this could be a limitation of the experimental design. If mothers find any specific emotional condition more difficult to 'act' and if this is related to a condition that was found to be associated with non-matching infant reactions, it could be that this is in fact influencing the infant's reaction.

Each of the eleven mothers was asked whether they found any of the emotional conditions more difficult to act, and they were asked to say why. They were asked to rate each condition on a scale from 1 to 6 (where 1 = very difficult and 6 = very easy). This means that the highest possible Total score (for a condition rated as very easy) would be 66, while the lowest possible Total score (for a condition rated as being very difficult to act), would be 11. Table 3.12 provides descriptive information for each condition.

Table 3.12: Mothers' rating of level of difficulty in 'acting' emotions

	<i>'Happy 1'</i>	<i>'Sad'</i>	<i>'Bored'</i>
Total	64	27	34
Minimum	5	1	1
Maximum	6	5	5
Mean	5.82	2.45	3.09
SD	.41	1.13	1.22

In general, mothers reported that they found that they did not really have to act the 'Happy' emotion -- they naturally felt happy as they interacted with their infants. It is of interest here that while mothers rated the 'Bored' condition as more difficult than the other emotions to act but, as the findings presented earlier show, it was an emotion that was easier for raters to 'Match'. Mothers also commented that they found the 'Sad' condition difficult to act. Mothers said that they felt it was unusual for them to be sad when in the company of their baby, without having any physical contact (cuddling or stroking). Others mentioned that if they felt sad in their daily lives, they

often tried to hide it from their infants. They found it difficult to be deliberately sad when talking to their baby.

Some mothers said that it was difficult to ‘act’ the emotion for the ‘Bored’ condition, simply because the infant was instantly less interested and less responsive to them. One mother said she felt “a bit silly” acting both ‘Sad’ and ‘Bored’ emotions, but another mother said she found the ‘Bored’ condition quite easy because she has two other children and is aware that she sometimes is bored in her role as a mother. Several mothers said that to make it easier to act in the ‘Sad’ and ‘Bored’ conditions, they had made facial expressions to highlight the emotion. In general mothers commented that it was easy to act the ‘Happy’ conditions because their babies were so responsive to them and some commented that they used their body more to show that they felt happy.

Table 3.13: Frequency information for emotional condition and mothers’ rating of difficulty

	<i>Observed</i>	<i>Expected</i>	<i>Residual</i>
‘Happy’	64	41.7	22.3
‘Sad’	27	41.7	-14.7
‘Bored’	34	41.7	-7.7

A Friedman test was carried out on the mothers’ total scores in each condition and found that there was a significant association between mothers rating of difficulty and the condition ($X^2=8.167$, $df=1$, $p=.004$). Table 3.13 provides details of the residuals for each condition which suggest that mothers did find the ‘Sad’ and ‘Bored’ conditions significantly more difficult to act.

To see if mothers’ level of difficulty in acting the emotions could have had an effect on the number of matches, a Spearman’s rho test was carried out. No significant association was found between the numbers of correct matches and the mothers’ rating of difficulty ($r_s=.120$, $N=33$, $p=.507$ ns).

3.4. Discussion and Conclusions

3.4.1: Emotional characteristics of the mothers' vocal expression in the three emotional conditions

Analysis of all mothers' vocalisations revealed that pitch and intensity differed significantly between the emotional conditions, which supports the hypothesis that these dimensions of vocalisations are important for expressing emotions differently. In particular, it was found that differences in pitch and intensity mainly distinguish between the positive and the two negative emotional conditions, and it may be that the mothers were using acoustic features of their voices to signal this difference in emotions. This corresponds to findings by Ladd et al (1985) who established that fundamental frequency (pitch) and intensity are the main acoustic cues used by adult listeners to judge affective state.

The findings presented here go beyond what was previously understood about where in the emissions of the voice these emotional cues lie. They suggest that single features of the mother's voice can provide consistent information about expressed emotions. A significant difference in pitch alone was found between the 'Happy' and 'Sad' conditions, between the 'Happy' and 'Bored' conditions, and between the 'Sad' and 'Bored' conditions. Evidently the pitch of vowel sounds is perceived as an important psycho-acoustic dimension in the expression of emotions and for perceptual differentiation of emotional states.

Results for intensity of vowel sounds showed a slightly different pattern. As for pitch, there were significant differences in intensity between the 'Happy' and 'Sad' conditions and between the 'Happy' and 'Bored' conditions, i.e. between positive and negative affect states. Unlike pitch however, the intensity of the mother's vowel sounds did not differentiate between the 'Sad' and 'Bored' conditions. This is different to the findings for all vocalisations, suggesting that vowels are playing a central role in the sharing of emotional meaning. Overall, it seems that mothers lower intensity in their vowel sounds when expressing negative emotions. This is consistent with findings presented by Scherer (1992) that showed intensity to be lower in adult

voices expressive of emotional states of 'boredom/indifference' and 'sadness/dejection'.

There was also a significant relationship between the frequency of vowels of different length and the three different emotional conditions. Mothers used more 'Long' and 'Extended' vowels when they spoke in a 'Sad' voice than when they expressed happiness or boredom. They used more 'Short' vowels when they used a 'Bored' voice than when they expressed happiness or sadness.

Wittmann and Pöppel (2000) define *tempo* as "the number of occurring events per time unit" (p. 3). They propose that *tempo* (and *rhythm*) allow the formation of expectancies, which in turn allow humans to use information from their senses to guide action. Wittmann and Pöppel place this process at the centre of interpersonal communication. In the present study, length of vocalisations appears to be an acoustic feature of vowel sounds that contributes to *tempo* or *rhythm*.

Donohue and Berg (1991) found that 7 month old infants quickly showed anticipatory responses (changes in heart-rate) to a temporal sequence including presentation of an interesting toy. The infants began to anticipate events and they could alter their expectations when the established sequences were experimentally disrupted. This shows that the tempo of events can assist infants to make predictions about what will happen and when, enabling them to adapt to unexpected changes.

It is known that temporal aspects of prosody such as duration, of sounds enable adults to differentiate between different types of dialogues in speech (Shriberg et al, 1998), or to predict turn-taking in conversation (Schegloff, 1998). The same has been shown for infants (Horner and Chetnik, 1986; Rochat et al, 1998; Vandenberg, 1999; Van Egeren et al, 2001). Crown et al (2002) examined the interactions of 45, 6-week-old infants with adults. They found that infants coordinated temporal aspects of their vocal and non-vocal communication with mothers and with strangers. Stern et al (1982) discuss the importance of this type of coordination of timing for

communication with infants, stating that “only when recurring patterns are recognisable can they become meaningful units of information” (p. 727).

Gratier (2006) describes how narrative sequences in musical expression are created “by variations of pitch and intensity, sequences of units of pulse and quality, that perform joint experiences of adventure and excitement” (p. 4). She goes on to explore how a similar process occurs in early infant communication. She suggests that systematic regulation of acoustic features in a mother’s voice helps the infant to begin to form expectations that support them to “situate and anticipate events and which they can adjust to fit their moment-to-moment communicative needs.” (p. 8). Trehub et al (1996) found that acoustic features of pitch (and tempo) were used by adult listeners to identify levels of emotional engagement, which they rated as higher in infant-directed singing. The findings presented in this study also suggest that emotion may be expressed and perceived by variations in pitch and intensity of vowels in communication with infants.

Infants sometimes vocalised concurrently with their mother in the ‘Sad’ and ‘Bored’ conditions, but not in the ‘Happy’ condition. The infants’ vocalisations were in the form of interruptions rather than the usual turn-taking vocalisations and were more often shouted than quietly spoken. Although this finding is limited because there were so few examples of this type of vocalisation, it may be that infants reacted in this way as a kind of protest when the mother expresses sadness or boredom.

3.4.2: Rating of infant non-vocal reactions to the mother’s voice

The hypotheses for this study predicted that infants’ gestures and expressions would ‘match’ or ‘mirror’ the emotion in their mother’s voice, and that this ‘matching’ would be discernable to adult raters. Results showed that there although raters made less ‘Matching’ than ‘Non-matching’ judgements, there was, indeed, a significant association between emotional condition and the frequency of ‘Matching’ responses. Overall for ‘Happy’ and ‘Bored’ conditions there were slightly more ‘Matching’ judgements made than for the ‘Sad’ condition. However, the difference was not

significant. There were significantly less 'Matching' than 'Non-matching' judgements in the 'Sad' condition than in the other conditions.

The lower number of 'Matches' for the 'Sad' condition is open to several interpretations. One possibility is that the infants may not have been convinced by their mothers 'acting' of this emotion. Another possibility is that indeed the infants did not react in a 'Sad' way when they heard this version of the rhyme. Some of the mothers reported that this emotion was more difficult to act, in particular without touching their infant. They said that in daily life, if they felt sad they were more likely to cuddle and touch their infants. It is interesting that some of the infants smiled in response to the mother's attempt to be sad and this was linked in some way to the mother's difficulty in acting the emotion.

Overall, the findings support the hypothesis that infants respond to specific emotions in a sympathetic way. In 'Happy' and 'Bored' conditions, infants were more likely to mirror the mother's affective state with a matching emotional gesture. One interpretation of this finding is that infants are sensitive to subtle emotional changes in the mothers' behaviour. It is interesting that the highest frequency (60%) of 'Matching' judgements were made for the 'Bored' condition. Infants' emotional reactions in this condition were avoidant and distressed and this is an important finding, relevant to anyone who has to communicate with infants when tired or withdrawn.

Infants were more likely to respond to their mother's 'Happy' and 'Bored' affective state in ways that could be clearly and consistently understood by blind raters who did not have access to any other contextual information, only the infants' gestural and expressive behaviour. There were differences in raters' abilities to read the emotion expressed by individual infants. Infants whose gestures and expressions were difficult to read tended to be more interested in (or distracted by) their surroundings and this made it more difficult for the raters to decide how the infant was reacting to their mother saying the rhyme. The emotions of some infants were more difficult to judge because they looked either very happy or very sad throughout all of the

conditions and their gestures and expressions did not seem to alter much during the test.

Papoušek (2007) found when using the 'still-face' paradigm that individual infants have different levels of ability to self-regulate. Different parenting styles and communicative environments contributed to the infant's response, but individual differences in the infant's ability to regulate arousal also played an important part. The results presented here add weight to the idea that individual infant temperament may reduce or increase an infant's ability to show a sympathetic response to emotions expressed in their mother's voice. This would be an interesting factor to consider in more depth.

3.4.3: Levels of difficulty in 'acting' different emotions

Mothers found the 'Sad' and 'Bored' conditions significantly more difficult to act, but there did not seem to be a relationship between mother's level of difficulty and number of 'Matches'. However, this aspect of the study raises a possible limitation to the significance of the findings presented here. It may be that infants can tell when mothers' expression of an emotion is not genuine or 'authentic'. This would correspond to Stern's (1985) finding that infants are able, even when very young, to deal with conflicting emotional information expressed through a mother's voice, face, gesture, position and proximity. As Stern says "attuning behaviour can be quite good even when your heart isn't in it. And as every parent knows, your heart can't always be in it, for all the obvious reasons from fatigue through competing agendas to external preoccupations that fluctuate from day to day" (p. 217).

The rationale for predicting that infants would mirror, in their posture, gestures, and/or facial and vocal expression, the emotion expressed in the mother's voice is based on two main theoretical perspectives. The theory of Innate Intersubjectivity proposes that features of movement transmitted by voice and gesture are both self-regulatory and goal directed for the infant, and that they provide "emotional definition" in signals to other persons (Trevarthen and Aitken, 1994, p. 599). Trevarthen (1993) interprets emotions as manifestations of 'motive states', and

describes how they are expressed and made evident to others by intricate coordination and timing of the changes of energy in the voice, hand gestures, and body movements. The theory of 'affect attunement' proposed by Stern (1985) proposes that communicating partners 'reflect' on another's affective states as they interact, and through this process of coordination expressions are validated, strengthened and built upon by each partner. Feldman (2006, p. 186) calls this process 'synchronous matching of internal cycles'. She found strong links between biological rhythms and the infant's ability to regulate its own affect and to participate in intersubjective communication.

There is much evidence to support these theoretical perspectives (Feldstein et al, 1993; Reddy et al, 1997; Fogel et al, 1999; Yale et al, 1999; Vandenberg, 1999; Van Egeren et al, 2001). Neuropsychological research leads to the conclusion that 'mirror neurons' in the brain react to vocal and gestural actions carried out by others, and by the self, and this equivalence of 'purpose' provides a mechanism for transmitting psychological meaning of physical actions between subjects (Kohler et al, 2002, Shibukawa et al, 2003). Panksepp and Bernatzky (2002) link perception of the 'melodic character' of acoustic information to physical aspects of brain processes, and neurochemical systems of the brain to various emotional experiences.

The evidence we have gained from the experiment with varied emotions in mothers' utterances and infants' reactions to them fits within these theoretical frameworks. It has been shown that measurable properties of mothers' speech, and in particular the properties of their vowel sounds, vary consistently when mothers express different emotions. In 'Happy' and 'Bored' emotional conditions infants were more likely to respond in characteristic ways to the emotional changes in their mothers' vowel sounds and this suggests that they are actively participating in a sharing of emotional meaning, and regulating their own affective state sympathetically.

Chapter 4: Can adults perceive emotions in infants' vowel sounds?

4.1: Introduction

From birth infants produce a range of vocalisations. It is possible that these early vocalisations may actively and consistently express different emotional states and that they may be reliably perceived as expressions of those emotions by the adults who are in communication with the infants. Unfortunately, we know little of what emotion these sounds may communicate. Much of the research that has been carried out to investigate what adults perceive in infant vocalisations has focused on infant cry sounds of distress, or has looked only for correlations with linguistic features without considering possible links with affective state. A few studies have looked at what experienced adults (i.e. parents and child care professionals) can consistently understand from the communicative sounds that infants produce.

Shimura and Imaizumi (1996) recorded vocalisations produced by one male infant each month between the ages of 6 and 17 months. They chose voice samples from each age to represent a variety of 'natural voice sounds' (p. 649). These samples were rated by students, nursery governesses and mothers who were still breast feeding their own babies. These adult listeners were randomly divided into two groups, and each group was given contextual information. One group was told that the vocalisations were produced in a situation when the infant was feeling pleasure, and the other group was told the vocalisations were produced when the infant was feeling discomfort.

The adults were then asked to rate each vocalisation against a list of 12 words that described 'aspects of emotion' ('happy', 'pleased', 'laughing', 'seeking affection', 'demanding', 'rejecting', 'frightened', 'sad', 'angry', 'speaking', 'singing' and 'crying'). The researchers carried out analysis of variance on rating scores to see if there were differences in the ratings for three factors; *listener group* (student, nursery governess or mother), *infant age* (6, 9, 12 and 17 months) and previously given *contextual information* ('pleasure' and 'discomfort'). They found that listeners could reliably

match with previous codings of 'pleasure' and 'discomfort', independent of infant age and previously given contextual information. However, matching previous coding of other 'aspects of emotion' was dependent on infant age. When they further analysed each 'aspect of emotion', they found that the listener group determined the level to which perception of emotion would depend on the contextual information that had been provided – mothers appeared to discriminate between more subtle aspects of the vocalisation rather than relying on the previously given contextual information. Students relied on contextual information and nursery governesses' rating scores fell between those of the other two groups.

Shimura et al (1996) continued this line of enquiry to see if younger children could match with adult listeners' coding of comfort or discomfort. First adult listeners were presented with a range of over 200 vocalisations produced by four 2-month-old infants. The listeners rated the vocalisations on a scale representing 'comfort' at one end and 'discomfort' at the other. The ten vocalisations with the highest and lowest ratings at each end of the scale were then presented randomly to two, three and four-year-old children. These children were shown two pictures of an infant's face - one representing 'comfort' and one representing 'discomfort', and they were asked to point to the picture that showed what they thought the emotional state of the infant was after hearing each vocalisation. They found that even young children aged between 2-3 years old could match the adult listeners' coding for the vocalisations. This research suggests that infants may be able to control their vocalisations to the extent that others will understand their affective state. However, it does not provide information about which particular aspects of the vocalisation were helping children and adults to perceive (or 'discriminate') the infants' emotional messages.

Papoušek (1992) reports a range of experiments that she and her colleagues carried out to examine acoustic correlates of emotional/behavioural information. Studies were carried out with German, American and Chinese mother-infant pairs. They then produced spectrographs of the 2-month-old infants' vocalisations to provide measures of frequency and amplitude. They carried out micro-analysis of corresponding video footage to categorise affective state ('pleasure', 'discomfort', 'cry' and 'joy') and found

that differences in some acoustic characteristics (including duration, amplitude, fundamental frequency and frequency range) were associated with particular affective states. This evidence led to the conclusion that “the physical structure of pre-syllabic interactional vocalisations reflects the infant’s behavioural-emotional states” (Papoušek, 1992, p. 238).

Papoušek and her colleagues then played a random selection of the infant vocalisations back to fathers and mothers of same age infants, primiparous and multiparous mothers of newborn infants, speech therapists and 8-year-old children. They found that all groups made judgements that matched the previously categorised emotional content of these single vocalisations, although there were differences between groups. Overall parents and speech therapists were significantly better than the 8-year-olds at matching with the previous categorisation. The 8-year-old children were better than adults at matching cry sounds and discomfort sounds and both fathers and mothers of same-age infants were the only group who could consistently match with previous coding of joy sounds. All groups made judgements about the level of intensity of the emotion they perceived, and it seems that judgements were often made in relation to the sound that listeners had heard previously. This may suggest that during natural interactions with infants, communicative partners are sensitive to varying intensity gradients between sounds and that they are able to estimate or detect moment-to-moment changes in affective state.

It seems likely that if adults can consistently judge the emotional content of infant vocalisations (including vowels sounds), then they are likely to have a sympathetic emotional response to what they hear. However, currently there is very little research to back up this suggestion. Papoušek (1991) found that American and Chinese mothers expressed “intuitive didactic care giving tendencies” (p. 243) in response to infant vocalisations coded as ‘neutral’, ‘comfort’, ‘joy’, ‘discomfort’ or ‘cry’. She divided mothers’ vocal responses into categories of ‘reward/greeting’, ‘encouraging a turn’, ‘encouraging imitation’, ‘evaluating infant state’, ‘reassuring of mothers’ presence’, ‘readiness to intervene’, ‘soothing’ or ‘discouraging’. Although there were some cultural differences, mothers tended to consistently differentiate their response

to match the emotion expressed in the infant vocalisation. Papoušek suggests that adults attribute meaning in response to the communicative sounds that infants make and that this has the evolutionary advantage of increasing the likelihood that the infant will be given security in parental care.

If adults can consistently sense the emotional content of the communicative sounds that infants make, this could suggest that infants have the ability to produce vocalisations that are adapted to provide consistent information about their psychological state. Single sounds were used in Papoušek's experiments, but people made judgements on relational aspects between sounds, so it may have been melodic information that people used to judge the emotional content. It seems likely, however, that there may also be some specific aspects of single vocalisations that provided affective information. There are many phonological or prosodic features that could be candidates for further analysis, but there are several reasons to hypothesise that extended 'vowel' sounds play a central role in the sharing of emotional information.

The first emotional vocalisations produced by infants can be described as 'extended vocables' (Locke, 2002). These early vocalisations contain many features that are similar to those used to produce canonical vowel sounds and it may be that they are the early precursors to formal vowel sounds. It is possible that acoustic variation in extended vocables, including identifiable vowel sounds, provides infants both with a means by which they can express their own feelings, and 'templates' to detect others' psychological states so they can begin to regulate the subtle personal relationships and learn the cultural meanings that are expressed and formed through vocal communication.

Extended vocables and vowels are known to use about 95% of the voice energy in their production (Lapp, 2004). Goldfield (2000) describes infants' exploration of acoustic vowel space as an emergent system – the *energy* (or *resonance*) of vowel sounds provides infants with a means by which they can “explore their own actions” (p. 433). The experience of energy variation of voice sounds plays an important part in a listener's perception of voice quality. Lapp (2004) describes voice quality (also

called ‘timbre’) as being one of the, “subtlest of all sound descriptors” (p. 11). He states that this aspect of voice gives an identity to the sound, or as he describes it “a flavour” (p. 11). Timbre involves complex acoustic phenomena and is identified as the interaction of several different acoustic features (such as the frequency of harmonics and the relationship between spectral aspects of these harmonics in a tone). It is difficult to tell which particular aspects of timbre are associated with the perception of differences in affective state (Tsang and Trainor, 2002). Nevertheless, it is known that differences in timbre play a major role in the expression and understanding of emotion in voices, as they do in the expression of music (Schogler, 2001; Trehub and Nakata, 2002; Tsang and Trainor, 2002; Malloch, 1997, 1999). Timbre certainly provides information about the intensity of expressed emotions (Johnstone et al, 2002), and as vowels and other extended vocables are known to contain a wide variety of timbral features, it is conceivable that these sounds in a mother’s production carry a considerable part of the information about emotional state to an infant listener.

It appears that parent-infant communication is also greatly facilitated by the mutual co-ordination of *rhythmical temporal patterning* of vowel sounds (Jaffe et al, 2001, Powers, 2001). There are innate timing and expressive mechanisms that facilitate the awareness of, and interactions with, other people (Stern, 1985; Trevarthen, 1999; Nagy and Monar, 2003). These mechanisms coordinate around motivational, sensory-motor and intersubjective systems and they operate within a cultural framework (Stern and Gibson, 1980; Stern, 1985; Stern et al, 1985; Stern, 1999). Theories of Communicative Musicality (Malloch, 1997, 1999; Trehub and Nakata, 2002) would suggest that acoustic features of infant and parent vocal interactions contain elements of pulse, quality and narrative that foster a sympathetic coordination of emotion and it may be that vowel sounds highlight emotional narratives and assist the coordination of temporal aspects of communication.

Consonants also define the temporal pattern in speech, and may play a more critical role in marking part of the tempo and rhythm of syllables than vowels do (Crystal, 1997; Lapp, 2004). Nevertheless, the evidence suggests that extended vocables in

general, and vowel sounds in particular, are marking the rhythmic pattern at another perhaps more basic level. It is likely that consonants, particularly in stress-timed languages, define the tempo of rapid articulations in speaking, and that extended vocables and vowel sounds provide additional information about when and how the energy of expression is changing over longer intervals of time during a communicative interaction. Vowel sounds can be taken to express the melody of emotion, and this melody provides communicative partners with a way of sharing information about one another's motives or intentions and feelings.

Ladd (1996) describes the extended harmonics (formants) of vowels as a salient feature of speech. Infants show a poorer performance on perceptual discrimination tasks with speech sounds when harmonic information is limited (Clarkson et al., 1996). It is possible that infants may hear vowel sounds as rich in perceptually attractive information because they are placed in a way that highlights the phonologically stressed aspect of a language (Fisher and Tokura, 1996). Sensitivity to vowels would correlate with the preference infants show for the affective content of infant-directed speech (Oller, 1986).

Linguistic reviews suggest that vowels play a dominant role in allowing the listener to gain contextual information about the next sound to be made and in intonation (Ladd, 1996; Kuhl et al, 1997). Vowels display information about the sex, tone, psychological state and age of the speaker and they give sequential and segmental details to the listener (Clarkson et al, 1996; Shimura et al, 1996). In this way they are relevant for affect and emotional expression and in the overall emotional narrative of any interaction as well as the articulation of more differentiated semantic information in speech. Vowel sounds can be seen as extended and modulated vocal 'gestures' and they have adaptive features for signalling and perceiving motives, emotions and intentions.

Research findings on the early development of speech suggest that vowels are prominent features that aid infants in the segmentation of speech and the perception of language specific rhythm class (Polka and Bohn, 1996; Sansavini, 1997; Davis and

Lindblom, 2001; Boersma et al, 2003). Kuhl et al (1997) found that infant directed speech in three cultures (Russian, American and Swedish) contained acoustically more intense vowel sounds and Kuhl and Meltzoff (1996) found that when infants hear particular vowel sounds they respond by imitating that particular vowel sound. So, vowels do appear to be acoustically attractive to infants, suggesting that they are components adapted to carry emotional information.

A great deal of research has been carried out to determine how emotional information is expressed and perceived by infants, and there is developmental linguistic evidence suggesting that acoustic features of vowel sounds are important for the mastery of speech, but there is little research to see if extended vocables of infants are primary carriers of emotional information.

The following experiment is designed to clarify the role of infant extended vocables or vowel sounds in the communication of emotion to adults by infants.

4.2: Method

Study 3 was carried out to answer the following research questions:

- Do infant vowel sounds convey specific, clear emotional messages to adults?
- How effective are females and males, and parents and non-parents at understanding emotional messages conveyed in infant vowel sounds?

4.2.1: Emotions in Infant Vocalisations – Stimuli

Isolated infant cries, calls and coos selected from the recordings of five English-speaking infants were chosen from the corpus of a previous study (carried out by Powers 2001 for MSc Thesis). Each vocalisation was an extended vocable resembling a vowel sound. Table 4.1 shows the age and sex of each infant, the sound that they produced, and the emotional content of the vocalisation (spectrographic information is provided for each vocalisation in Appendix 1).

Vocalisations were rated for their emotional content within the overall communicative interaction observed from video and described according to the category system of Trevarthen and Marwick (1981) for definition of 'affect' (see Appendix 2). Each of the ten interactions was 1 minute in length and was analysed in 5 second segments. Raters used both acoustic and visual aspects of the video to make their decision about the emotional content of each mother-infant interaction from moment to moment. The coding of each interaction was tested for inter-rater reliability by 2 raters, one of whom was blind to the purpose of the experiment. Reliability was established at 93%. One isolated vowel sound was then taken from each of the ten recordings to represent a 'positive' and a 'negative' vocalisation (for each of five infants) according to the preceding ratings, and these were used as the acoustic stimuli for the current test. In Study 3 presented here all codings from (1) to (9) of the Trevarthen and Marwick scale were taken to be 'Positive' and codings from (10) to (23) were taken to be 'Negative' (See Appendix 2).

Table 4.1: Infant Age, Sex, Description of Vocalisations and Emotional Message

Sound File (Infant sex and ID)	Infant age in weeks	Vowel Sound	Emotional Content
1. (Male 1)	23 weeks	/i/ (shrieked)	Positive
2. (Male 1)	23 weeks	/u/	Negative
3. (Female 2)	23 weeks	/ʌ/	Positive
4. (Female 2)	23 weeks	/æ/ (shouted)	Negative
5. (Female 3)	28 weeks	/awuε/	Positive
6. (Female 3)	28 weeks	/æεu/	Negative
7. (Male 4)	36 weeks	/uε/	Negative
8. (Male 4)	36 weeks	/æ/	Positive
9. (Female 5)	46 weeks	/yæ/	Positive
10. (Female 5)	46 weeks	/aw/	Negative

4.2.2: Participants

The study was run via the Internet. 158 English-speaking participants were recruited from Edinburgh University, Lothian NHS Board and by word of mouth. They were asked to rate the infant vocalisations for emotional content. They also completed a short questionnaire to provide information about their sex, age and status as a parent or non-parent. Participants who lived with children who were not their biological relation were defined as parents if they had lived with the child(ren) for more than 6 months.

Table 4.2 provides details of participants' age, sex and parental status. Their ages ranged from 18 to 59 years. The 'non-parents' were much younger than the 'parents'. There was also fewer 'male parents'.

Table 4.2: Participant sex and parental status

<i>Sex</i>	<i>Parental Status</i>	<i>Number</i>	<i>Mean Age (yrs)</i>
Female	Non-parent	59	21.12
Female	Parent	36	42.08
Male	Non-parent	45	23.82
Male	Parent	18	43.75

The ten vocalisations were presented in a random order for each participant. Each participant was asked to rate each vocalisation as either 'Positive' or 'Negative', or as 'None' (containing no emotional message). Definitions of 'Positive' and 'Negative' were provided and participants were able to listen to the vocalisation as many times as they needed to before making their decision. 'Positive' vocalisations were defined as expressions that convey the feelings of 'happiness', 'joy' or 'pleasure' and 'Negative' vocalisations were defined by any emotional expression considered to be 'sad' or 'unhappy'. After judging each sound for emotion, participants were asked to provide additional written details of the emotion that they thought they had heard in the sound. They were then asked to provide written details of any emotion they had felt in response to the infant vocalisation.

4.2.3: Analysis

Judgements about emotion

In order to assess whether adult participants in the different groups could perceive specific emotions in infant vowel sounds, the number of times their judgements matched previous rating of the emotional content of infant vocalisations (as judged by raters of the videotaped interaction) was counted and totalled for each group. An overall one-tailed prediction was made that adult participants would judge infant vocalisations as containing some kind of emotion. Different participant groups -- female or male and parent or non-parent -- were compared to see if there was any association between sex or parental status and ability to correctly judge emotional content of infant vowel sounds. It was predicted that participants would be able to match emotions to the infant vocalisations but that there would be differences in this ability between females and males, and between parents and non-parents.

Written Descriptions of emotion in infant vocalisations

As well as being asked to judge what emotion was contained in the infant vocalisations, participants were also asked to use their own words to describe any more detail that they thought that they perceived about the emotional message that they felt was conveyed in the vocalisation. They were told that they could use any words to describe their thoughts or feelings about the sound that they heard. This was designed to explore what information participants wished to add, based on the small sample of acoustic information that they had heard. Comparisons of female with male, and parents with non-parents were carried out.

Written Responses to emotion in infant vocalisations

Participants were asked if they felt any emotional response to the infant vocalisations that they heard, and these responses were categorised. Analysis was carried out according to a two-tailed prediction to see if there were differences between the types of response felt by different participant groups. A further analysis, again with a two-tailed prediction, was made to see if there were any associations between response type and 'Matching' or 'Non-matching' judgement type.

4.3: Results

This experiment was intended to investigate whether measurable physical aspects of infants' calls resembling vowels, produced by continuous muscular changes during phonation, can create distinctive acoustic signals that express motive processes and convey emotional messages that can be consistently recognised by adults.

4.3.1: Perceiving emotions in infant vowel sounds

Participants' judgements of emotional expression in relation to infants' vowel sounds were categorised and counted. Participant responses were categorised into 'Matching' answers (i.e. matching with the researcher's previous coding for emotional content and inter rater coding), 'Non-matching' answers (i.e. perceiving an emotion that did not match with the previous coding), and 'None' where no emotion was perceived in the infants' vowel sounds.

The dataset consisted of 1,580 judgements (from 158 participants each listening to 10 sound files). For 1,100 of these judgements, participants reported that they perceived emotion (either positive or negative). Chi square analysis showed that when 'Matching' vs. 'Non-Matching' scores are combined and compared with 'None' there was a significant difference from chance ($X^2 = 1580.000$, $df = 1$, $p < 0.001$). So, overall participants did perceive infant vowel sounds as containing some kind of emotional message.

4.3.2: Judgements about emotion

'Matching', 'Non-matching' and 'None' responses were counted giving the following totals:

- 874 'Matching' judgements,
- 218 'Non-matching' judgements,
- 498 'None' judgements.

A chi square analysis showed that overall participants could 'Match' with previous codings of emotion ($X^2 = 488.442$, $df = 2$, $p = .0001$).

Analysis was carried out to compare females/males and parents/non-parents. Table 4.3 shows frequencies and percentages for each type of response for each participant group and it appears there was a similar pattern of judgement types for all participant groups.

Table 4.3: Descriptive information for sex, parental status and judgement type

Parent	Sex	Judgement Type	Frequency			Percentage	
			Observed	Predicted	Pearson Residual	Observed	Predicted
Non-parent	Female	Matching	202	211.578	-1.006	54.6%	57.2%
		Non-matching	53	49.428	.546	14.3%	13.4%
		None	115	108.994	.685	31.1%	29.5%
	Male	Matching	115	105.422	1.398	60.5%	55.5%
		Non-matching	29	32.572	-.688	15.3%	17.1%
		None	46	52.006	-.977	24.2%	27.4%
Parent	Female	Matching	321	311.422	.806	56.3%	54.6%
		Non-matching	63	66.572	-.466	11.1%	11.7%
		None	186	192.006	-.532	32.6%	33.7%
	Male	Matching	236	245.578	-.895	51.3%	53.4%
		Non-matching	73	69.428	.465	15.9%	15.1%
		None	151	144.994	.603	32.8%	31.5%

The residuals shown in Table 4.3 suggest that 'female non-parents' were more likely (than other groups) to make 'Matching' as opposed to 'Non-matching' and 'None' judgments. 'Female parents' and 'male parents' were more likely to make 'Matching' judgments as opposed to 'Non-matching' and 'None' judgements and male non-parents were more likely to make 'None' judgments.

To explore this data in more depth, judgement types were examined by creating a Multinomial Logit Stepwise (MLS) Model to compare associations between judgement type ('Matching', 'Non-matching' and 'None' as outcome measures) and participant group -- 'female non-parents', 'female parents', 'male non-parents' and 'male parents' -- as predictor values. This type of regression analysis was used as it provides more power than a chi square analysis, it can show main effects and stepwise effects as variables are entered into the model and it allows a more general comparison for a large number of categorical variables. Assumptions for this model are that any two response patterns are independent from all others and that any effects

will have a proportionately equal affect on all factors. Sex and parental status were put into the model to see if they would be predictive of judgement type and responses were compared by using ‘Non-matching’ as the reference category. Table 4.4 provides a summary of the analysis.

Table 4.4: Summary of MLS analysis of participant group as a predictor of judgement type

<i>Predictor variable for judgement type</i>	<i>β (regression coefficients)</i>	<i>Std Error</i>	<i>df</i>	<i>Sig</i>
‘Matching’				
Female	.280	.153	1	.068
‘None’				
Female	.323	.165	1	.050

The table shows that females are more likely than males to make ‘Matching’ judgments ($\beta = .280$, $p = .068$) although this difference is not significant. Females are significantly more likely than males to make ‘None’ judgements ($\beta = .323$, $p = .050$). Parental status was not found to be predictive of judgement type.

4.3.3: Effect of Sound files

Pitch information for all sound files containing infant vowel sounds is shown in Appendix 1. The vowel sounds were taken from both negative and positive emotional interactions. Infant age and sex also varied.

Table 4.5 (next page) shows the number of judgements of each type for sound file and all participant groups combined.

Table 4.5: Frequency Information for Sound Files by Judgment Type and Participant Group

Judgment Type	Participant Group	Sound File (+ = Positive, - =Negative)										Totals
		1 (+)	2 (-)	3 (+)	4 (-)	5 (+)	6 (-)	7 (+)	8 (-)	9 (+)	10 (-)	
Matching	female parents	12	21	19	31	21	15	27	21	22	13	202
	female non-parents	33	39	28	31	25	31	32	36	34	32	321
	male parents	7	13	11	17	10	11	11	13	15	7	115
	male non-parents	21	22	19	27	26	24	27	22	24	24	236
	Totals	73	95	77	106	82	81	97	92	95	76	874
Non-matching	female parents	10	4	4	1	12	3	2	8	3	6	53
	female non-parents	10	6	8	5	9	7	5	5	3	5	63
	male parents	9	1	2	0	4	1	3	2	2	5	29
	male non-parents	10	10	7	7	6	6	1	8	8	10	73
	Totals	39	21	21	13	31	17	11	23	16	26	218
None	female parents	15	12	14	5	4	19	8	8	12	18	115
	female non-parents	14	12	21	21	23	19	20	16	20	20	186
	male parents	3	5	6	2	5	7	5	4	2	7	46
	male non-parents	15	14	20	12	14	16	18	16	14	12	151
	Totals	47	43	61	40	46	61	51	44	48	57	498

To explore these patterns in more depth ‘sound file’ was added into the MLS model and compared to see if there was a difference between ‘Matching’, ‘Non-matching’ (which was used as the reference category) and ‘None’. Table 4.7 (next page) provides summary information of the findings.

Table 4.7: Summary of MLS analysis of sound file, as predictors of ‘Matching’ judgement type

<i>Predictor Variable for ‘Matching’ judgements</i>	<i>β (regression coefficients)</i>	<i>Std Error</i>	<i>df</i>	<i>Sig</i>
File 4	1.027	.372	1	.006
File 7	1.106	.391	1	.005
File 9	.710	.353	1	.045

The table shows that File 4 ($\beta = 1.027$, $p = .006$), File 7 ($\beta = 1.106$, $p = .005$) and File 9 ($\beta = .710$, $p = .045$) do have an effect on participants’ ability to ‘Match’ with previous judgements of emotion suggesting that some sound files are easier for raters to match with emotion.

4.3.4: ‘Non-Matching’ Judgements

Participants made some emotional judgements that did not match with the researcher’s previous coding and these ‘mistakes’ were analysed to see if there was any difference in the types of mistakes made. Table 4.8 provides information about how participant groups judged the vocalisation when they did not ‘Match’ in relation to the initial coding.

Table 4.8 ‘Non-Matching’ Judgements

	<i>Number of sound files not matched with previous coding, incorrectly judged to be...</i>	
Participant Group	A ‘Negative’ emotion	A ‘Positive’ emotion
Female non-parents	32	33
Female parents	21	32
Male non-parents	39	34
Male parents	9	20

The table suggests that there is a fairly similar pattern of 'Non-matching' responses for all participant groups, although it is interesting to note that 'parents', male and female, make fewer 'Non-matching' 'Negative' mistakes than 'Positive'.

Chi square analyses were carried out to see if there was a significant association between participant group and mistake type. A 2 x 2 chi square showed that there was no significant association between sex and mistake type ($X^2 = 0.101$, $df = 1$, $p = 0.750$ ns). However a separate 2 x 2 chi square found a significant association between parental status and mistake type ($X^2 = 4.576$, $df = 1$, $p = 0.032$) indicating that parents are more likely to make a 'Non-matching' judgement that the vocalisation contains a 'Positive' emotional message. In other words when they get it wrong, they often perceived the vocalisation as 'Positive'.

4.3.5: How do adult participants describe infant emotions?

As well as being asked to judge what emotion was contained in the infant vocalisations (i.e. in terms of 'Positive', 'Negative' or 'None'), participants were also asked to use their own words to describe in more detail what they thought about the emotional message, or what they felt was conveyed in the vocalisation. They were told that they could use any words to describe their thoughts or feelings about the sound that they heard.

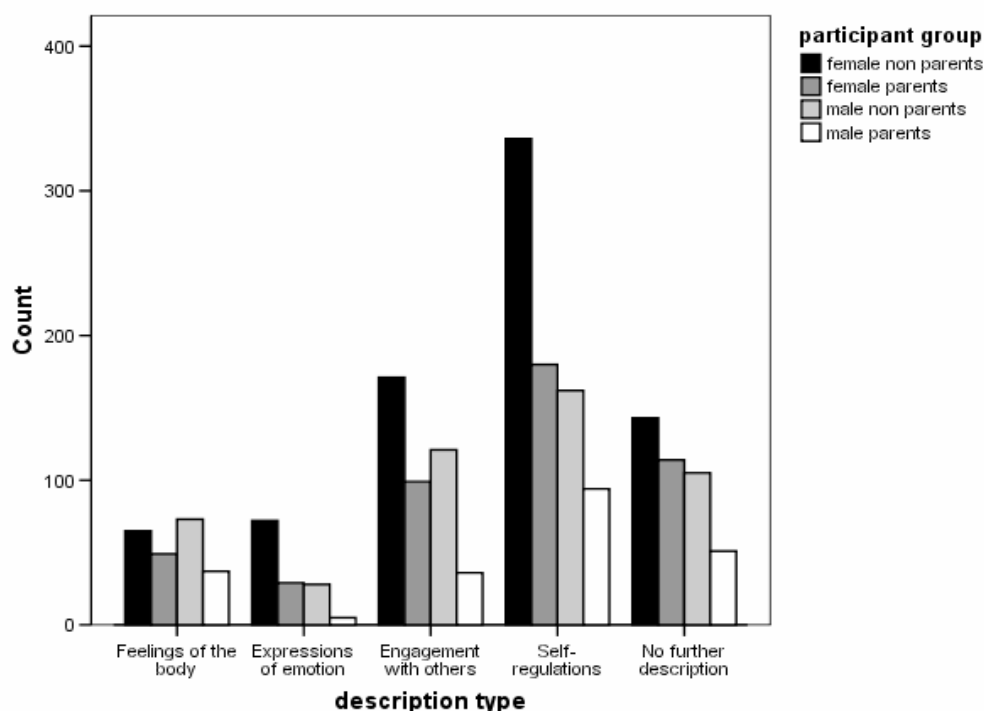
The written descriptions given by participants after they had listened to each file were collated, and descriptions from 'Matching', 'Non-matching' and 'None' responses were analysed. To simplify analyses of these verbal descriptions, all words associated with emotion were coded according to the following categories:

1. 'Feelings of the body': for example – hunger, pain, tiredness, discomfort.
2. 'Expressions of emotion': for example – laughter, crying, about to cry, thinking about laughing, starting to cry.
3. 'Engagement with others': for example – playing, being tickled, upset at something, seeking communication, greetings, playing with voice, acknowledgment, recognition, needing attention.

4. 'Self-regulating': for example – interested, showing off (pride), distress, relaxed, restless, playful, enjoyment, wonder. This category also included descriptions that reinforced the participant's original judgement of the sound as 'Positive' or 'Negative', for example happy or sad.

Individual participants wrote different verbal descriptions of the emotions that they perceived in the infant vocalisations. Some wrote single words, whereas others wrote phrases or sentences and analysis was carried out at the word level. Some verbal descriptions (single words and phrases) contained more than one of the above categorisations. The vowel sound segments were very short and there were 6 comments from participants complaining that some vowel sounds were too short for them to make a judgement. These comments were not included in the analysis. Descriptions that did not make sense were excluded from all analyses. Details of these findings are presented in Graph 4.1:

Graph 4.1: Description Type for Females and Males



The graph suggests that there are some differences in the types of description given by each participant group. Female non-parents give more 'Self-regulation' descriptions of the emotion that they perceive in the infant vowel sounds than the

other groups. Male parents give more ‘Self-regulation’ descriptions than they do other types of description.

A main-effects MLS Model was constructed with description type, participant group and judgement type to explore the effect of each variable. Table 4.6 provides a summary of findings from this analysis.

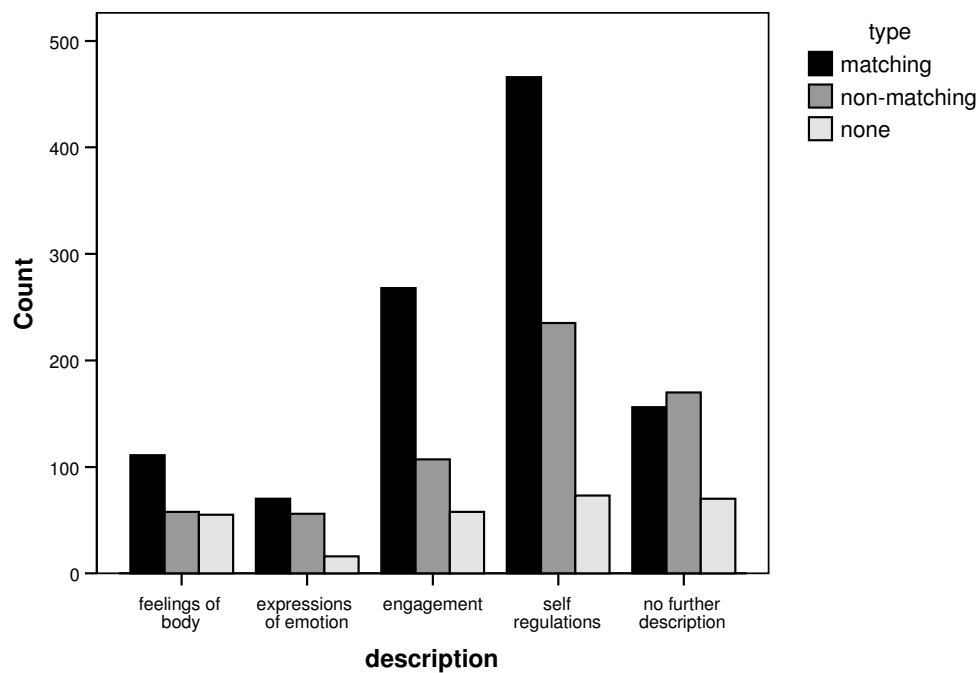
Table 4.6: MLS Summary for Description Type

<i>Variable (Type of description)</i>	<i>β (regression coefficients)</i>	<i>Std Error</i>	<i>df</i>	<i>Sig</i>
Females				
‘Feelings of the body’	.522	.179	1	.004
‘Expressions of emotion’	.423	.210	1	.044
‘Engagement with others’	.692	.150	1	.0001
‘Self-regulations’	.481	.136	1	.0001

The analysis shows that sex is predictive for description type. Females are significantly more likely than males to give each type of description. Parental status showed no significant relationship to description type.

Graph 4.2 (next page) shows that the types of description given are different when participants are ‘Matching’ with previous judgements as opposed to when they are ‘Non-matching’, or when they say that there is no emotion contained in the sound. The graph shows that when participants make ‘Matching’ judgements, they give a higher frequency of ‘Engagement with others’ descriptions and ‘Self-regulation’ descriptions than other types of description, to further explain the emotion that they perceive in the infant vowel sounds.

Graph 4.2: Judgement Type by Description Type



To explore the influence of variables in the data, a main effects MLS model was constructed, and Table 4.7 provides a summary of the main findings of this analysis.

Table 4.7: MLS Summary for Judgement Type and Description Type

<i>Variable (Type of description for judgement type)</i>	β <i>(regression coefficients)</i>	<i>Std Error</i>	<i>df</i>	<i>Sig</i>
‘Matching’				
‘Feelings of the body’	.735	.196	1	.0001
‘Engagement with others’	1.004	.159	1	.0001
‘Self-regulations’	.771	.137	1	.0001
‘None’				
‘Feelings of the body’	.834	.236	1	.0001

The table confirms the descriptive findings. When participants make ‘Matching’ judgements they are significantly more likely to give ‘Feelings of the body’ descriptions ($\beta = .735$, $p = .0001$), ‘Engagement with others’ descriptions ($\beta = 1.004$, $p = .0001$) and ‘Self-regulations’ descriptions ($\beta = .771$, $p = .0001$).

An unexpected result of the analysis of verbal descriptions was that participants who judged the infant vocalisations as containing 'no emotion' still went on to provide written descriptions of the infant's state, often giving detailed emotional descriptions, contradicting their 'None' judgement. The forced choice reduced the amount of information they were inclined to offer. However, they were significantly more likely to give 'feelings of the body' descriptions ($\beta = .834$, $p = .0001$).

4.3.6: Adults' emotional responses to infant vocalisations

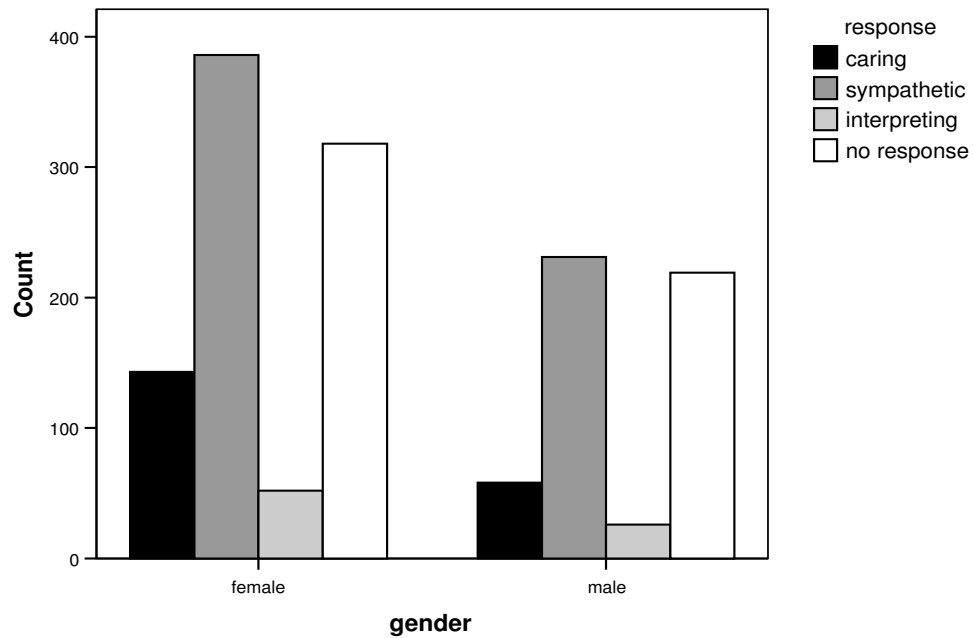
Adult participants were asked if they felt any emotional response to the infant vocalisations that they heard and their responses were categorised according to the following descriptions:

1. 'Caring': a need or desire to take action, for example – to soothe, comfort, pick up, attend to, pull faces / make sounds for the infant.
2. 'Sympathetic': a direct emotional response, for example – concern, happiness, curiosity, pleasure.
3. 'Interpreting': a conceptual response extending the verbal description to cover other factors such as the environment or assumptions about the stage of development or the infant's actions or intentions, for example – 'infant changing her boundaries', 'the baby is desperate to succeed', 'the infant requires something but I don't know what', 'could imagine the baby being in my arms and looking up at me'.

Any written responses that did not make sense were excluded from all analyses, and participants' emotional responses sometimes contained words from several of the categories.

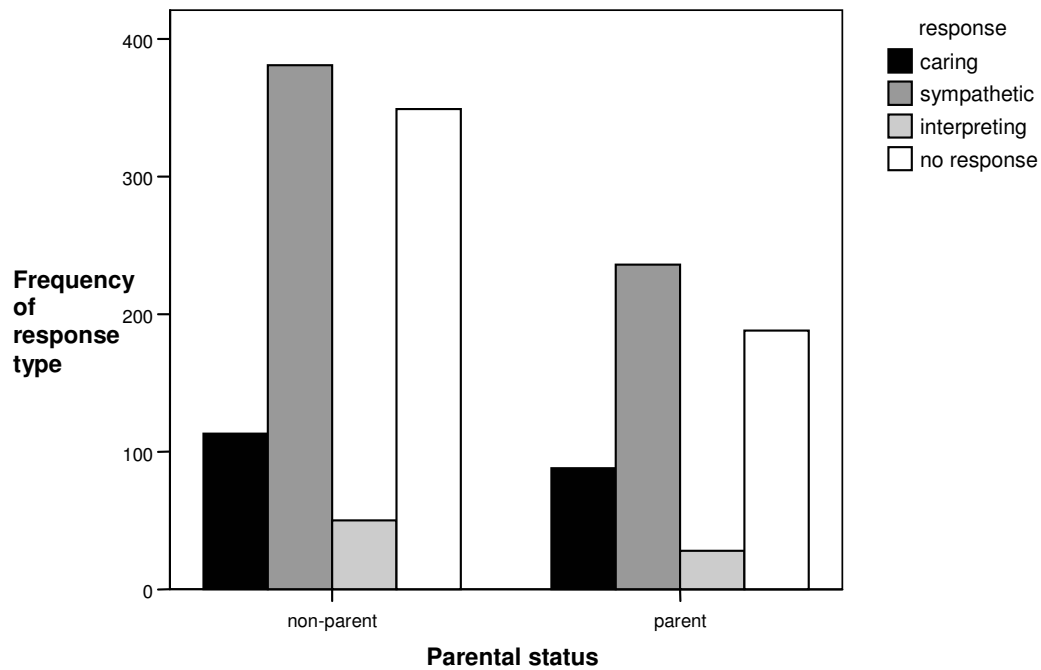
Graph 4.3. (next page) provides information about response type for 'females' and 'males'. It shows that 'males' exhibit less 'Caring' responses to the infant vocalisation than 'female' participants. Conversely 'females' show more 'Sympathetic' responses than 'males'.

Graph 4.3: Response Type for Females and Males



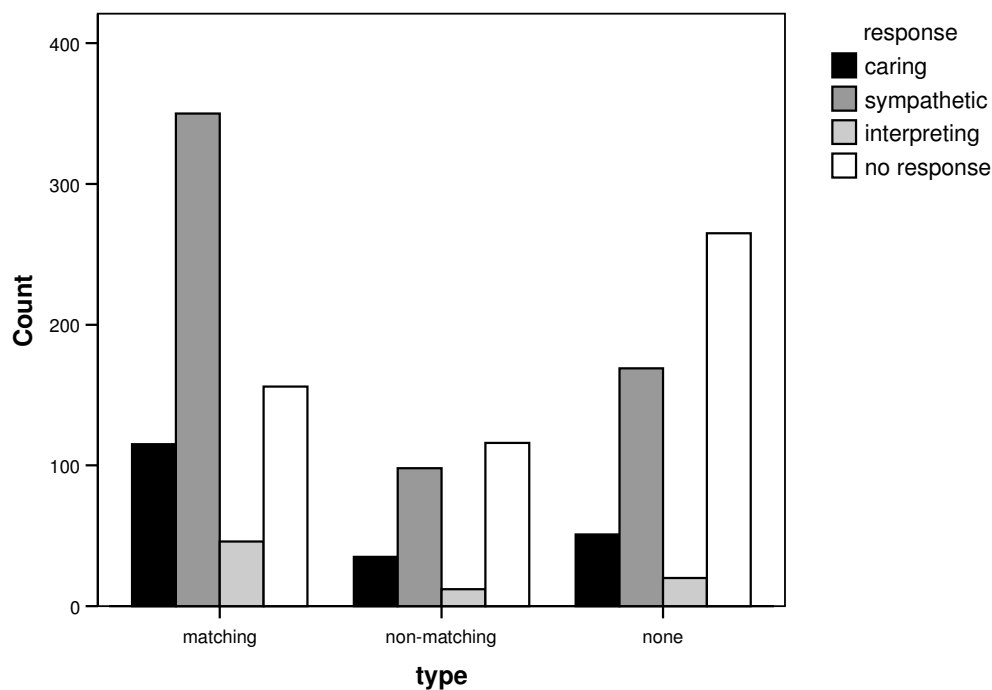
Graph 4.4 provides information about response type for 'parents' and 'non-parents'. It shows that in general 'non-parents' felt more of each type of emotional response to the sound that they heard than 'parents' did.

Graph 4.4: Response Type for Parents and Non-parents



A main effects MLS model was created to explore whether sex and parental status was predictive of particular responses. Female sex was found to be associated with a 'Caring' response ($\beta = .529$, $p = .003$) and being a 'parent' was associated with feeling a 'Caring' response ($\beta = .369$, $p = .029$).

Graph 4.5: Response Type by Judgement Type



Graph 4.5 shows that for the 'None' respondents there was a higher number of participants saying that they felt 'No response' to the infant vocalisations. Again a MLS model was calculated to explore these associations further. Participants were significantly more likely to give 'Caring' responses ($\beta = .893$, $p = .0001$), 'Sympathetic' responses ($\beta = .977$, $p = .0001$) and 'Interpreting' responses ($\beta = 1.047$, $p = .003$), than when they made 'Matching' judgements.

4.4: Discussion and Conclusions

4.4.1: Understanding emotions from infant vowel sounds

As expected, adults could 'Match' with previous codings of emotion, suggesting that they did perceive emotional messages in isolated vowels sounds. This shows that infants can effectively express their emotions in vowel sounds alone in a way that can be consistently understood by adults.

It was predicted that participants would be able to match previous coding of the emotional content of the infant vocalisations and that 'parents' would be able to do this

more consistently than 'non-parents'. The results confirmed this prediction. 'Females' were also more likely than 'males' to make 'Matching' judgments of the emotion in infant vowel sounds, although this difference was not significant. Parental status was not predictive of judgement type. In relation to sex this corresponds with research by Shimura and Imaizumi (1996) who found that mothers compared to fathers were able to discriminate between more subtle aspects of emotion in infant vocalisations.

There is a considerable amount of information available about the vocalisations that infants produce before they can speak. For example acoustic aspects of infant vocal communication are consistently used by 10-month-old infants to differentiate 'communicative' vocal behaviour from 'investigative' vocal behaviours (Papaeliou and Trevarthen, 2006). Infants (from 2 to 14 weeks old) produce different types of vocalisations depending on their parents' vocal and non-vocal responses to them (Keller and Schölmerich, 1987; Masataka, 1992) and as they develop infants increasingly coordinate acoustic aspects of their vocal and non-vocal behaviour (Malatesta, 1981; Hsu et al, 2001) to produce ever more complex vocal melodic (intonation) contours (Hsu et al, 2000).

Research suggests that melodic contours of infants' vocalisations are used in vocal communication to express specific emotions, and that they are regulated by coordination of acoustic and gestural movements of communication (Ferland and Mendelson, 1989; Papoušek et al, 1990; Papoušek et al, 1991). Papaeliou et al. (2002) used pitch pattern recognition software to explore pitch patterns in vocalisations produced by 7 to 11 month old infants as they communicated with their mothers. They found that infant vocalisations categorised by mothers as being 'emotional' (as opposed to being 'communicative' or 'informative') used longer *duration*, and had lower *pitch* and *intensity*. This suggests that infants are using acoustic features of their voice to express emotion. The findings presented here provide more detail concerning whether emotion in infant vocalisations can be consistently understood by adults, including those who are not their parents.

It seems that some infant vowel sounds are significantly predictive of parents' ability to make 'Matching' judgements for emotion. These files contain vowel sounds with both 'negative' and 'positive' emotional messages and they are produced by male and female infants of different ages, so it seems likely that these factors do not influence a listener's ability to judge emotion. It is not possible in the present study to say with certainty why these particular files are matched more often than the others. Some of the vowel sounds selected for this test were more obviously conveying one emotion or another and, from listening to the files, it seems to be the less expressive or more neutral files that are difficult to match. This would be a variable requiring future research to examine a wider range of vowel sounds, and produced by a larger number of infants.

The judgements of the adults did not always match with the previously made codings of the emotional content of infants' sounds. 'Non-matching' responses were similar across all participant groups, but it is interesting to note that there was a significant association between parental status and mistake type indicating that 'parents' are more likely to make a 'Non-matching' judgement when the vocalisation contains a 'Positive' emotional message.

Parents are known to use a combination of vocal and non-vocal information to interpret their infant's affective state (Green et al, 1995). Some research has been carried out to explore what emotional signals parents perceive using pictures of infant facial expression. Spangler et al (2005) showed pictures of infants expressions identified as 'positive', 'negative' and 'neutral' to parents, and measured parent response to these pictures on three levels: by asking for their affective judgements; by attaching facial electrodes to measure mimic/sympathetic response; and by measuring the physiological startle response to collect information about autonomic reactions. They found that parents could make judgements that matched with previously coded categories of emotion, and that they mimicked the emotion that they perceived in the pictures, but parents also showed a bias for 'positive' interpretation of infant affect. The researchers suggest that this has an adaptive advantage. If parents perceive a negative emotion but they respond with positive affect, they are more likely to initiate a caring response.

4.4.2: Adult emotional descriptions of and responses to infant emotional voices

The results presented here show that females and males, parents and non-parents describe infant vocal affect in vowel sounds, in different ways. 'Female non-parents' were significantly less likely than other participant groups to give 'Feelings of the body' descriptions, than they were to give 'social' or 'emotional' descriptions. 'Non-parents', female and male, were significantly more likely to give 'Expressions of emotion' descriptions and 'Engagement with others' descriptions to describe the emotion they perceived in the infant vocalisations. When participants made 'Matching' judgements they were more likely to give 'Engagement with others' and 'Self-regulation' descriptions. It is known that parents respond differently and attribute different emotional meaning to different infant emotional vocal and facial expressions (Keller and Schölmerich, 1987). Carvajal and Iglesias (2006) looked at adults' emotional judgements of infants with and without Down's syndrome. They found that adult judges, who were not parents and who did not have experience with infants, or people with learning difficulties, could match a previous categorisation of infant emotion based on an emotion recognition coding system. The researchers tested both vocal expression and facial expression of infants with and without Down's syndrome. They found that adult judges could more easily match with previous judgements for infants by facial expressions, and for infants who did not have Down's syndrome.

Izard et al. (1980) found that undergraduate students and female health service professionals could reliably match a previous emotional categorisation by a facial movement coding system of 1- to 9-month-old infant expressions of the so-called 'basic' emotions of 'interest', 'joy', 'surprise', 'sadness', 'anger', 'disgust', 'contempt' and 'fear'. Emde et al (1985) found similar results when they asked adult participants to respond to photographs of infant faces. Camras et al (2002) found that adults could consistently identify surprise and interest, using a range of non-verbal gestures and movements in European-American, Japanese and Chinese infants aged 11 months.

The present study goes beyond these findings to provide more detail about how infant vocal affect is understood by adults and how parental status and sex affect this perception. It also shows how adults respond emotionally to infant emotions expressed in vowel sounds. The results interestingly showed that 'male non-parents' were significantly less likely to feel a 'Caring' response to infant emotions when compared to the other participant groups. However, the results do not reveal why this might be.

Tronick (1989) suggested that parents and infants constitute an 'emotional communication system' and the infant's communicative behaviour and the corresponding response of parents operating together have a direct impact on the infant's emotional well-being and development. Indeed the results of the current study could be interpreted from such a socio-emotional perspective. It could be speculated that, because they take care of infants in the first year of life, parents become more attuned to providing care and so are more likely to feel a caring response to infant vocalisations. With regard to sex, and from an evolutionary perspective, it may be that because females can give birth, they have some innate mechanism or social 'expectancy' that makes them more likely to feel a caring response or have had experience of caring.

When the participants of the present study made 'Matching' judgements, they were significantly more likely to feel 'Caring' responses to the infant vocalisation. When they made 'Non-matching' or 'None' judgements they were significantly more likely to feel a 'Sympathetic' response or an 'Interpreting' response to the infant vocalisations. This suggests some link to affective attunement as discussed in previous chapters. If adults are 'in tune' with infants they are more likely to be caring, if they are not so 'in tune', they may want to try to understand the infants' affective state, or they may feel sympathetic towards them in a more rational way.

Reddy and Trevarthen (2004) propose that emotions are "an important agent in an infant's active, moving and assertive relationship with the world" (p. 10). They suggest that infant emotions provoke emotional 'sympathetic' responses from those

who interact with them. The findings presented here support this idea and indicate that adults were more likely to make ‘Matching’ judgements of infant emotion as opposed to ‘Non-matching’ and ‘None’ judgements. One interpretation of this is that infant vowel sounds can convey affective state in ways that can be consistently understood, and appropriately responded to, by others.

Chapter 5: General Discussion and Conclusions

The research presented in this thesis was undertaken to explore expressions of emotion in the voices of mothers and infants, and their perception. Attention was restricted to the extended and modulated sounds of vowels, or vocables comparable with vowels. These were hypothesised to be a principal medium of social signalling by means of which emotional meaning is shared.

Vowels sounds of various kinds, emitted in a range of situations were recorded and submitted to detailed acoustic analysis. Experiments were made to test infants' ability to discriminate adult emotions from speech, and to respond in sympathetic expressive ways to them, and to prove the ability of men and women, parents and non-parents, to perceive emotions in infants' sounds. A comparison was made of vocal communication of emotion in Japan and Scotland, two countries with very different speech and different theories of human emotions and how they should be expressed in society.

The idea that physical movements of behaviour can be studied to explain psychological experience of emotion is not a new one. In 1872, Darwin proposed three guiding principles for expression of emotions that made explicit links between physical actions and states of mind. These principles state how purposeful actions are linked with states of mind. Darwin believed that actions serving immediate needs of the body are repeated until they become habit and then the resulting habits become reflex actions associated with experiences, and in this way become in animal society a vehicle for communication of inner mental events. Since Darwin's pioneering work, which included a rich account of the expression of emotions by infants, researchers have been trying to establish how and which physical acts of communication are related to inner mental experience in those who make or perceive the expressive movements.

The research presented here chose to analyse the variety of vowel sounds, extended vocables or calls of defined length, that appear to be adapted for adult-infant communication, to see if they are linked to particular 'states of mind'. Such a reductionist approach, seeking to measure the features of only one isolated element of a complex process, carries the risk of oversimplification. Nevertheless, such a reduction, or focus of enquiry, is often a necessary first step toward obtaining a meaningful insight into a part of any natural process that is not yet fully understood. The functional 'usefulness' of vowel sounds as a channel through which physical events of action and interaction are translated into the psychological experience of emotion, is certainly not fully understood.

Combinations of vocal and non-vocal movements of communication with infants have been described many times on the assumption that they are linked to psychological experiences (Stern et al, 1982; Beebe et al, 1982; Beebe and Gerstman, 1984; Scherer, 1992; Reddy et al, 1997; Donald 2001), but little is known about the functions of vowel sounds in particular. Kuhl (1994) has proposed that vowel sounds heard by infants serve as perceptual magnets, which pull attention and perceptual learning towards a few locations in the map of vocal sounds, and it may be that those sounds made by other persons' vocal tracts that attract infant attention are also related to the motor coordination processes by which infants are able to communicate and express their motivational and affective states.

The studies presented here aimed to provide detailed description of pitch, intensity and duration of vowel sounds produced by mothers and infants in two different cultural contexts. The first hypothesis was the following: if acoustic features of vowel sounds exchanged between mothers and infants are coordinated and used differently, and consistently, to express distinct affective states in a human community, then they have been proved to provide a medium through which emotional meaning can be shared, even with very young and culturally inexperienced human beings. Secondly, the extremely different cultural contexts of Japan and Scotland were explored to find whether acoustic features of vowel sounds are used by mothers differently but systematically in culturally specific ways which might affect

the infants' developing expressive system. Producing a description of vowel sounds produced during naturally occurring mother-infant interactions provided a framework within which to go on and test the idea that mothers' and infants' vowel sounds can be consistently understood by the infants and by other adults, or to reveal the effects of cultural ritualisation of vocal communication.

In Study 1 relationships were found among pitch, intensity and duration of vowel sounds in mothers' and infants' voices. These relationships were found on two levels – within individual mother-infant pairs and on a cultural level, in Scotland and Japan. This relationship was also found when vowel sounds were isolated and analysed. Acoustic features of vowel sounds -- pitch, intensity and duration -- were varied in systematically different ways by both mothers and infants during periods of 'engaged' and 'disengaged' interaction, and in different types of tactile play, although the consistency of relationships between acoustic features was less clear for infants. It was found that at 4 months old infants are showing the initial signs of using acoustic features of their vowel sounds in coordinated and consistent ways. The pitch of mothers' vowel sounds increased when they tried to encourage their infants, and mothers increased the intensity of their vowels during active play. Infants increased the intensity of their vowel sounds when their mother was exercising their legs and arms in playful ways.

These findings are important, because they provide further insight into how infant-directed speech is structured to facilitate communication with infants. It is known that infant-directed speech has predictable expressive patterns that are based on *voice quality*, *pitch*, *intensity* and *melody* (Malloch et al, 1997). Infants can discriminate pitch information of simple sounds much as adults do, although unfamiliarity with tonal patterns or distorted sounds can cause difficulty for infants (Cohen et al, 1987; Clarkson et al, 1996). Infants' perception is tuned in to the enhanced features of adult expression when the adults are addressing the infant.

In Study 1, 4 month old infants were found to be starting to combine acoustic features in their vowel sounds in ways that corresponded with combinations created by the mother. The way that the mothers and infants were using acoustic features of their

vowel sounds was explored here during periods of spontaneous rhythmical play. The findings suggest that there is mutual purposefulness in the coordination of acoustic features of the voices in these engagements. The voice is used 'in action', and it seems that vowel sounds provide one way in which mothers and infants can have awareness of each other's shifting emotional states by sympathetic detection of the movements that express them.

Trevarthen (1993) describes the process of communication with infants as "an active mutual engagement of highly specific intersubjective motives or programs of communicative expression in mother and infant" (p. 145). He describes an array of vocal and non-vocal media for the sharing of "mental dynamics", one of which is the controlled variation of pitch, intensity and timing in the voice. The findings presented here show that these variables are prominent in mother and infant vowel sounds. To explore in more detail the ways in which adults and infants understand the emotional content of vowel sounds, two further studies were carried out.

The results of Study 2 showed that mothers' pitch range and intensity differed significantly for vowel sounds when they tried to express different emotions while reciting the same nursery rhyme. Mothers increased the pitch range of their vowel sounds when expressing 'Happy' emotions and they increased the intensity of their sounds when expressing negative emotions of 'Sadness' and 'Boredom'. This confirms that pitch and intensity in vocalisations are important components for expressing specific emotions and for differentiating between positive and negative emotions. Differences were found in the distribution of vowel lengths in different emotional conditions. Mothers used more 'Long' and 'Extended' vowels when they spoke in a 'Sad' voice than when they expressed themselves as 'Happy' or 'Bored'. They used more 'Short' vowels when they used a 'Bored' voice. This consistent coordination and differentiation of acoustic (physical) features of mothers' vowel sounds seemed to facilitate expressive and perceptual differentiation of emotional (psychological) states.

Individual infant temperament appeared to reduce or increase an infant's ability to show a sympathetic response to emotions expressed in their mother's voice, but overall infants responded consistently to systematic changes in pitch, intensity and duration of mothers' vowel sounds. Infants' gestures and expressions were analysed to see if they were judged as matching or mirroring the emotion in their mother's voice when she spoke the nursery rhyme. The majority of 'Matching' judgements, showing agreement between the mother's assumed emotion and raters' judgements of the infants' feelings observed on video, were made for the 'Bored' condition and infants' emotional reactions in this condition were judged to be 'avoidant' and 'distressed'. Infant gestures and expressions were more likely to be judged as 'Non-matching' for recordings made when the mother was reciting the verse in the 'Sad' condition. The level of difficulty mothers reported in acting the different emotions was not found to be related to 'Matching' judgements in the above test. Infants vocalised concurrently with their mother more often in the 'Sad' and 'Bored' conditions, whereas they rarely vocalised at all in the 'Happy' condition.

It seems that physical properties of mothers' speech, and in particular their vowel sounds, are consistent when mothers seek to express specific emotions. Raters judged that infants responded in characteristic ways to emotional changes in their mothers' vowel sounds. The findings support the theory that infants between the ages of 3 and 9 months respond to specific emotions in a sympathetic way, that they are sensitive to subtle emotional changes in the mothers' vocal and gestural actions, and that this provides a mechanism for them to represent psychological meaning of physical actions and to reinforce their own ability to regulate emotional interactions.

Study 3 was carried out to explore whether infant vowel sounds can convey specific, clear emotional messages to adults, and to see if contrasting groups of adults, females vs. males and parents vs. non-parents, understand emotional messages conveyed in infant vowel sounds in the same or different ways. Results showed that adults could often 'Match' previous coding of specific emotions from tiny samples of infant vowel sounds, and that the sex of the adult was predictive of their ability to match with previous codings of the emotional content in the infant vowel sounds. 'Females' were

more likely than 'males' to give 'Matching' judgements as opposed to 'Non-matching' judgements. Interestingly parental status was not predictive of adults' ability to match with previous judgements of emotion in infants' vowel sounds. These must be accepted as preliminary findings require to be validated by replication. However, they provide an intriguing insight into how the sex of an adult, or perhaps their socially shaped 'gender', affects perception of infant emotion. The differences observed could be interpreted as showing that females having some biological propensity towards perceptions that aid infant care-giving, while, at the same time, parents' abilities to recognise infant emotions are, in part, learned, or affected by other factors which were not tested in this study.

Previous studies have found that many factors may affect parental sensitivity to infant emotions. Pesonen et al. (2004) considered parental attachment relationships and parental self-rated depressive symptoms to explore how these affected parental judgements of infant temperament. They found that parents who displayed depressive symptoms made more negative judgements about their infant's temperament. Pelchat et al. (2003) asked parents to recall their own experience of being parented and measured levels of parental stress, stress within their relationship and socioeconomic information. They then analysed these factors to understand how they were related to parental sensitivity, for infants with Down's syndrome, for infants with cleft lip or palate and for normally developing infants. They found that a mother's sensitivity to infant emotion was related to her levels of education and family income, while a father's sensitivity was related to their own experience of being parented, levels of stress in their relationship, income level and their infant's condition. Kaitz and Maytal (2005) found a relationship between mothers' anxiety level and levels of disruptive mother-infant communication. It may be concluded that many factors can affect parents' perception and response to emotions expressed by infants, and this perhaps highlights a limitation of the current study. In future research it would be desirable to include measurement of aspects of adult personality, emotional intelligence and on their views and perceptions of infants and care giving practice and experience.

An interesting issue raised by this study is why three sound files of vowel sounds produced by three different infants were predictive of adult ability to match with previous codings of the infants' emotions based on more complete information from video. This leads to the conclusion that some vocalisations made by infants are more easily understood by other persons. Two of the sound files contained 'Negative' vocalisations by female infants, and one contained a 'Positive' vocalisation by a male infant. This is insufficient information on which to base a firm conclusion about why some infant expression can be more easily understood by others.

The findings presented show that the sex of the adult can be predictive of the type and richness of description given by adults of infant vocal expressions and of emotional judgements made about infant vowel sounds. Females were more likely than males to give all types of further description when they expanded on their initial judgement of emotion contained in the infant vowel sounds. Their descriptions of infants' feelings were broadly classified as 'Feelings of the body', 'Expressions of emotion', 'Engagement with others', or 'Self-regulating'. Females were more likely than males to give all description types although parental status was not associated with description type.

The examination of adults' emotional responses to infant vowel sounds revealed that females were more likely than males to feel a 'caring' response to the emotional sound that they heard, and parents were more likely than non-parents to feel a caring response. When adults made 'Matching' judgements they were more likely to have caring, 'sympathetic' and 'interpreting' responses, as opposed to 'no emotional response'. 'Matching' in this study implies that the adult making the judgement is more sensitive to infant emotional expression because they have concurred with previous judgements, which were made with the benefit of full non-verbal and contextual information.

Schoppe-Sullivan et al (2006) examined the roles of infant and parent sex and attachment in relation to parental sensitivity to 1 year-old infant emotions. They found that overall fathers were less emotionally sensitive to daughters than mothers

were, and that mothers were less emotionally sensitive to sons. Within families they found that fathers were particularly sensitive to sons who had insecure attachment relationships with their mother. Zeifman and College (2003) showed parents video clips of 4 week-old infants fussing or crying. They found that 'dimensions' of parents' personality ('extraversion', 'empathy' and 'low levels of conscientiousness') were associated with parental sensitivity to expressions of infant emotion. They also found a strong link between parental child-rearing attitudes and perception and response to infant crying. It may be concluded that many factors appear to influence adults' response to infant emotion, and although the study presented here does not address these factors, the findings presented here are important because they address questions that have previously not been asked. They show that adults perceive specific emotions in infant vocalisations (even in a vowel sound of less than a second long, and even although they sometimes can't always 'Match' a given emotion with its previous coding).

In conclusion, this research has answered some important questions in terms of how acoustic features of mother and infant vowel sounds are coordinated for expression of emotion and sharing of emotional states. The findings begin to fill some of the gaps in our understanding of how features of the voice contain predictable patterns that allow us to create meaning in our engagements with one another and to adapt to new conventions of communication.

There, of course, are limitations in the findings presented here. It would have been better to have repeated the experimental aspects of this research in Japan as well as Scotland. Findings would have been improved if Japanese coders had also coded emotional engagement and acoustic features of vowel sounds. It remains to test Japanese adults' perceptions and responses to infant emotions expressed in vowel sounds and to code Japanese infants' reactions to variations in emotional expression of their mothers' voices.

Nevertheless, the research has triggered some interesting reflections. It described how acoustic features of vowel sounds are used during 'engaged' and 'disengaged'

communication, and in 'positive' and 'negative' affective states. We are led to a number of clear questions. These are examples:

Are acoustic features of vowel sounds coordinated in consistent ways when more subtle emotional states than those investigated here are expressed?

Do infants respond in consistent ways to a broader range of emotions expressed by their mothers, including emotions that are less self-related than 'happiness', 'sadness', and 'boredom'?

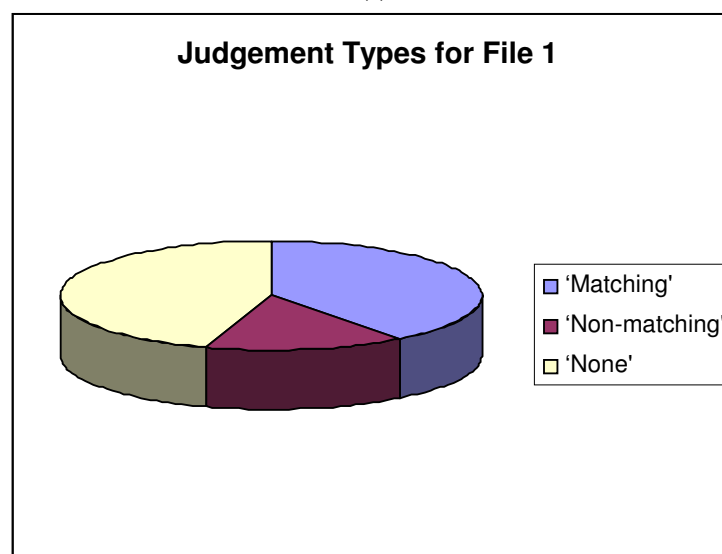
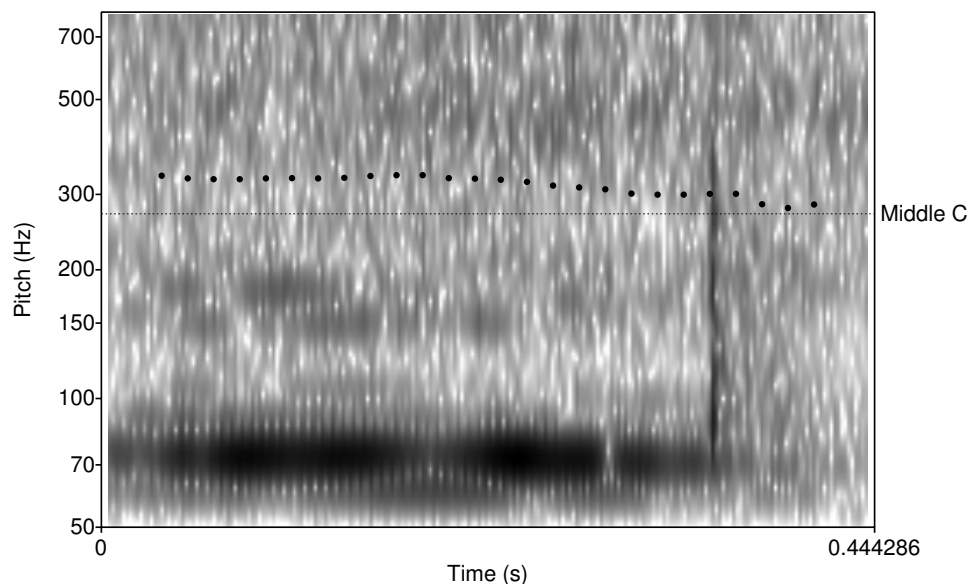
Does the same process occur when fathers, other adults or young children communicate vocally with infants?

Future research will aim to address some of these questions.

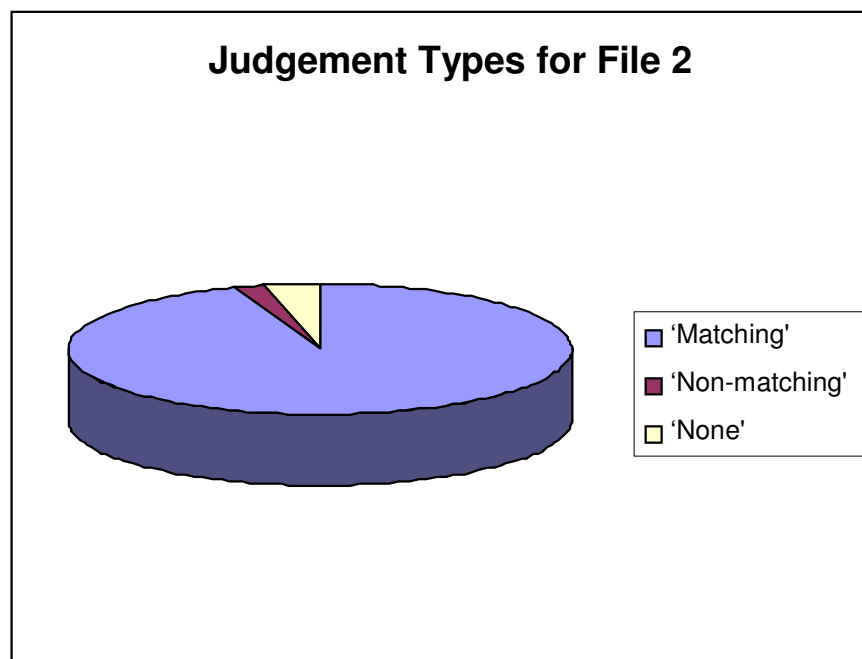
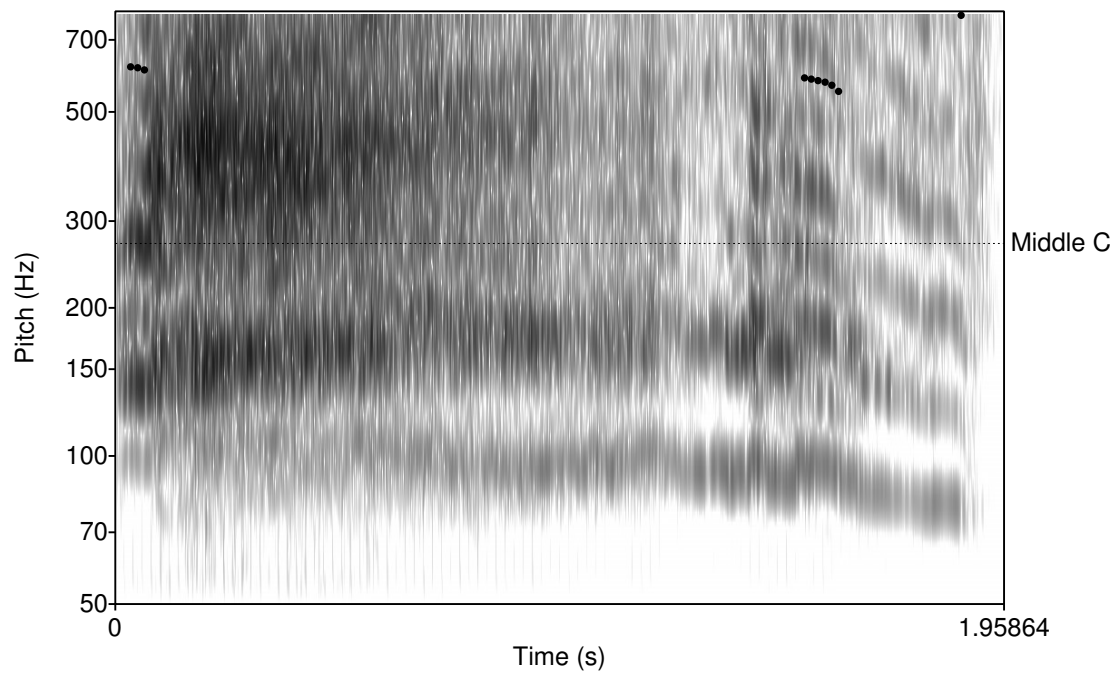
Appendix 1: Pitch plots of sound files

All files are different lengths (PRAAT analysis stretches, or shortens the sound automatically to fit the graph). The dotted line shows the pitch analysis of the sound. Most of the files are of pre-linguistic vocalisations sometimes produced before the infant is able to articulate canonical vowels and consequently pitch information is sometimes sparse. Dots closely spaced together show richer pitch information. Spectrographic information is shown in shades from light to dark grey (representing Hertz) and from black to white (representing Decibels).

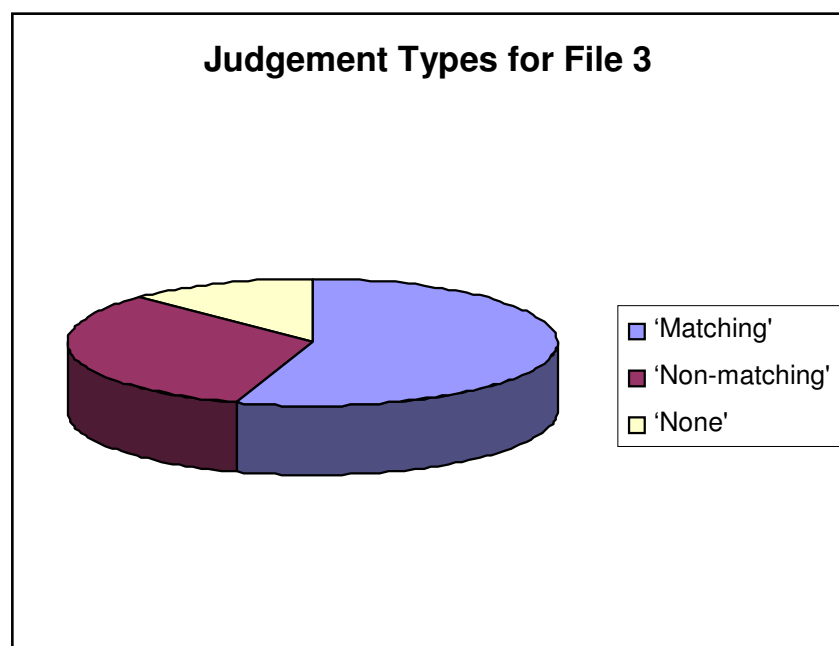
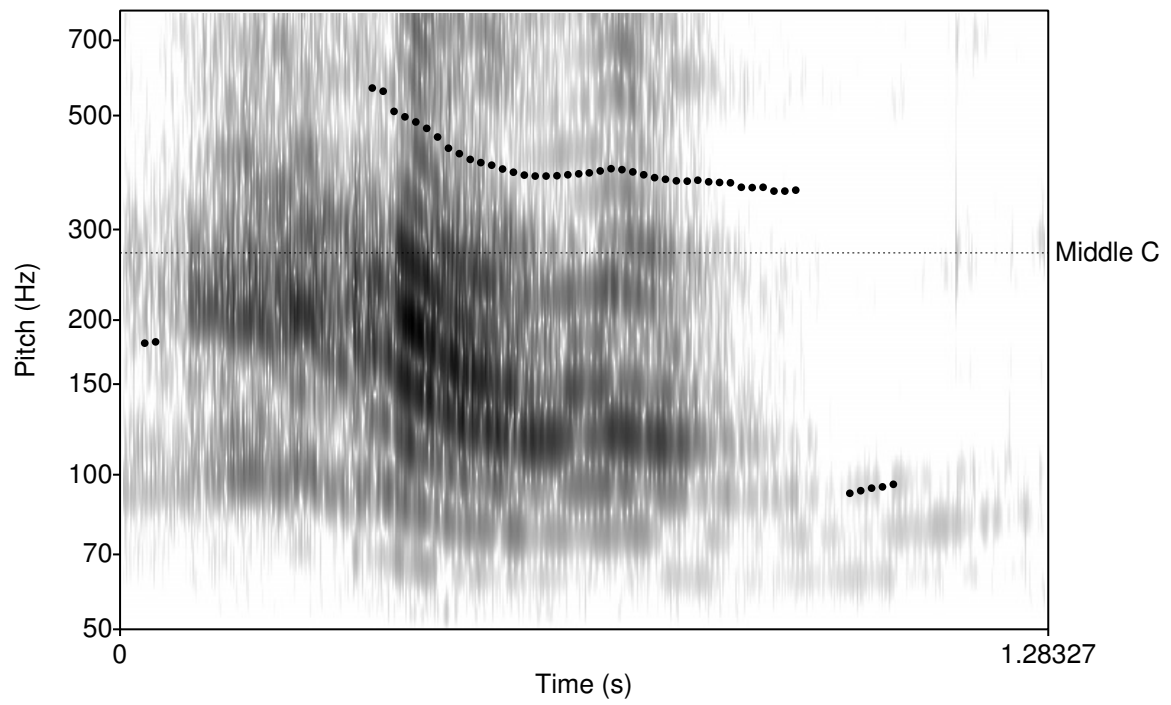
File 1: Ho1, 23 week old female infant, positive emotion ('uh' sound)



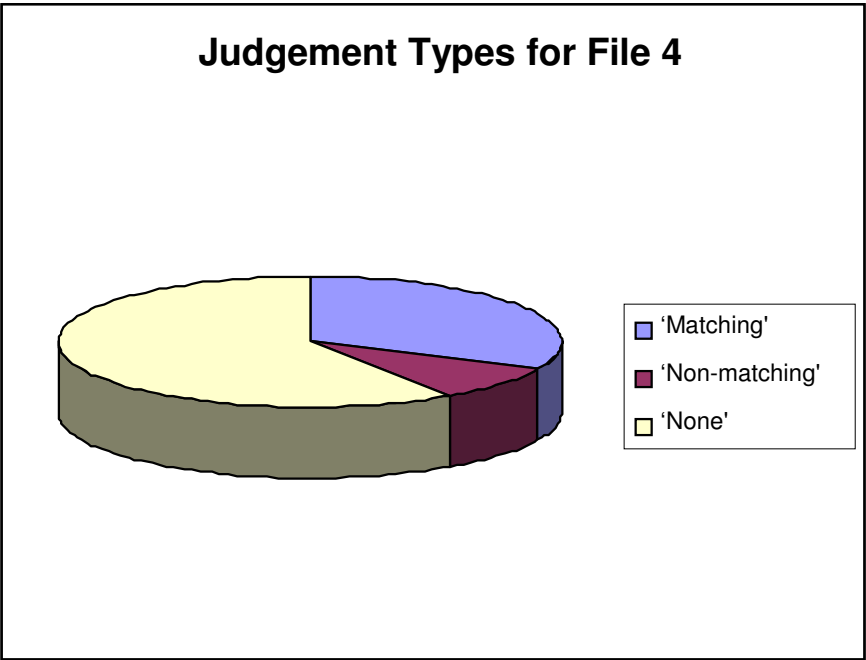
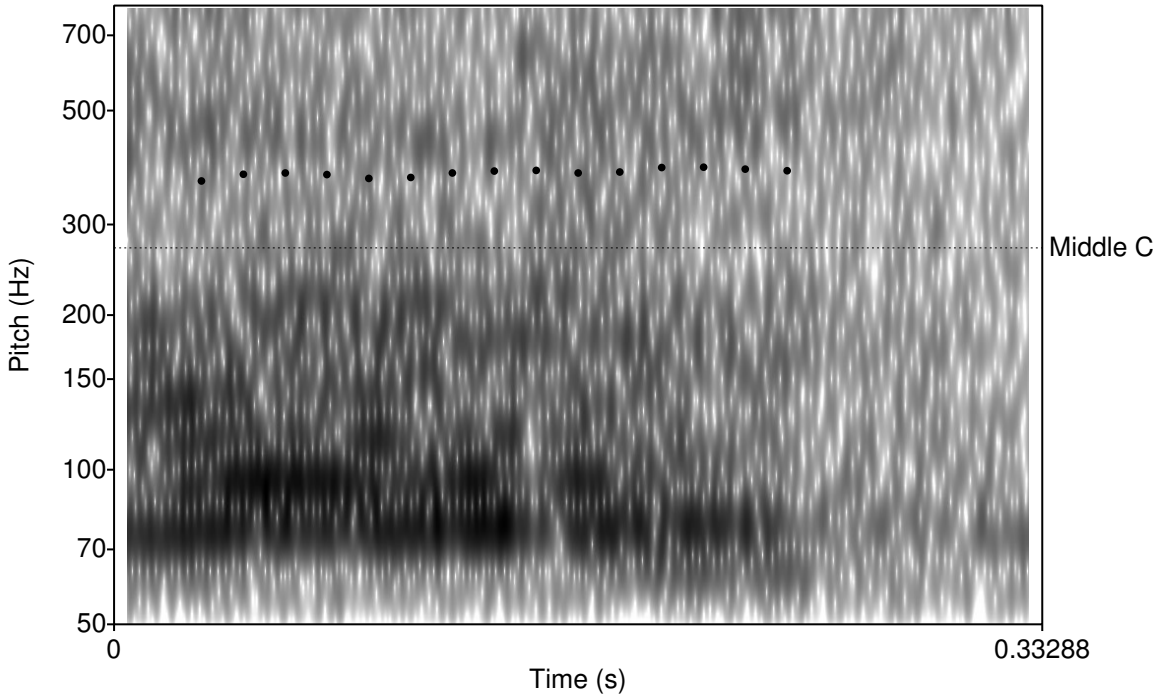
File 2: Ho2, 23 week old, female infant, negative emotion (A shrieked 'a' sound – scratchy sound so pitch couldn't be measured completely)



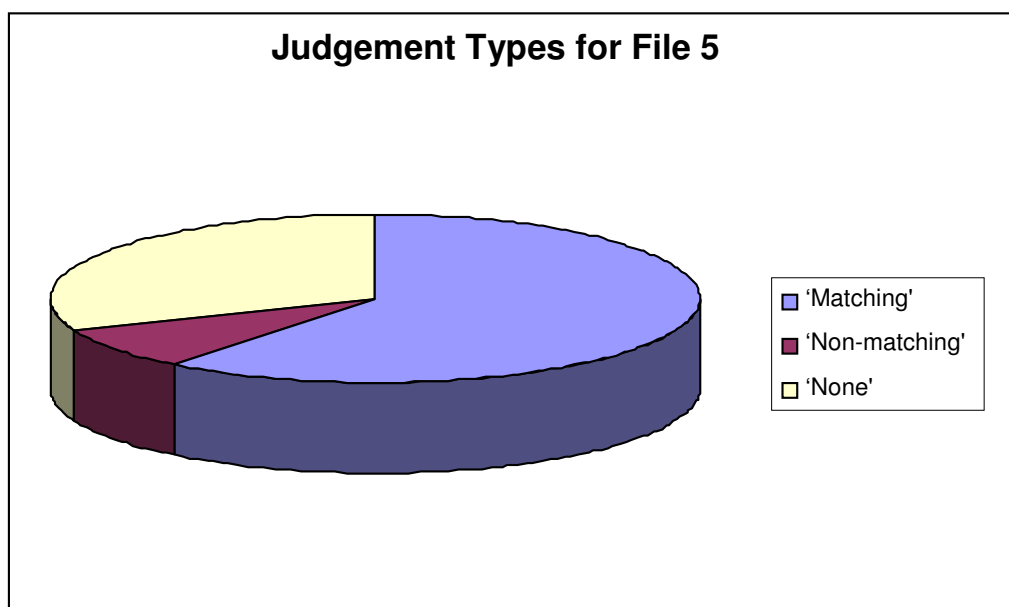
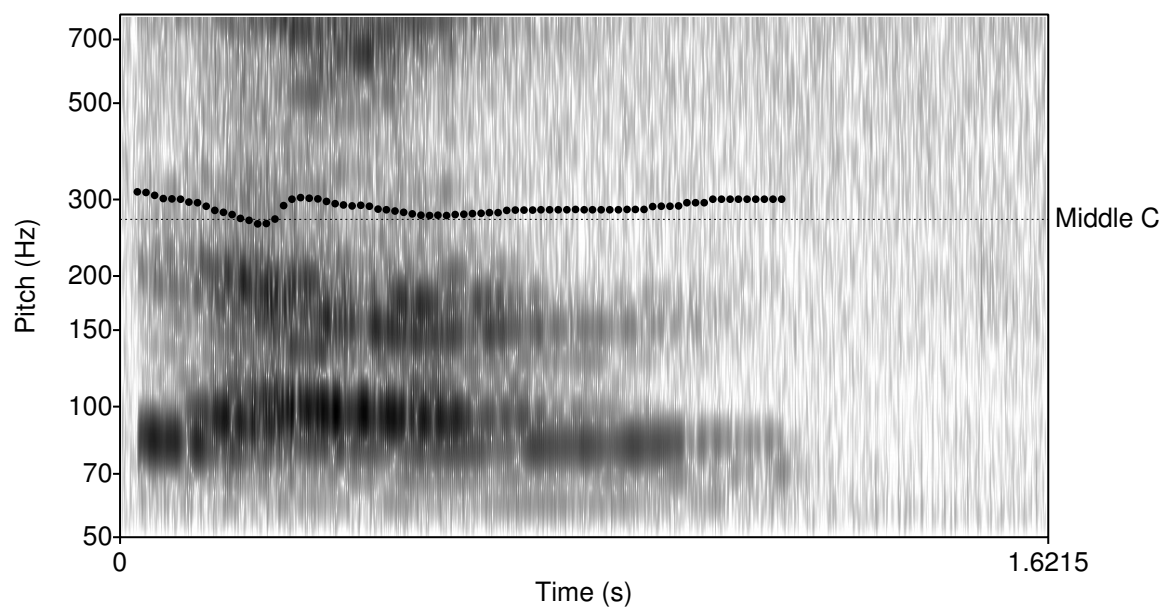
File 3: Is1, 23 week old, male infant, positive emotion (A shrieked 'ee' sound)



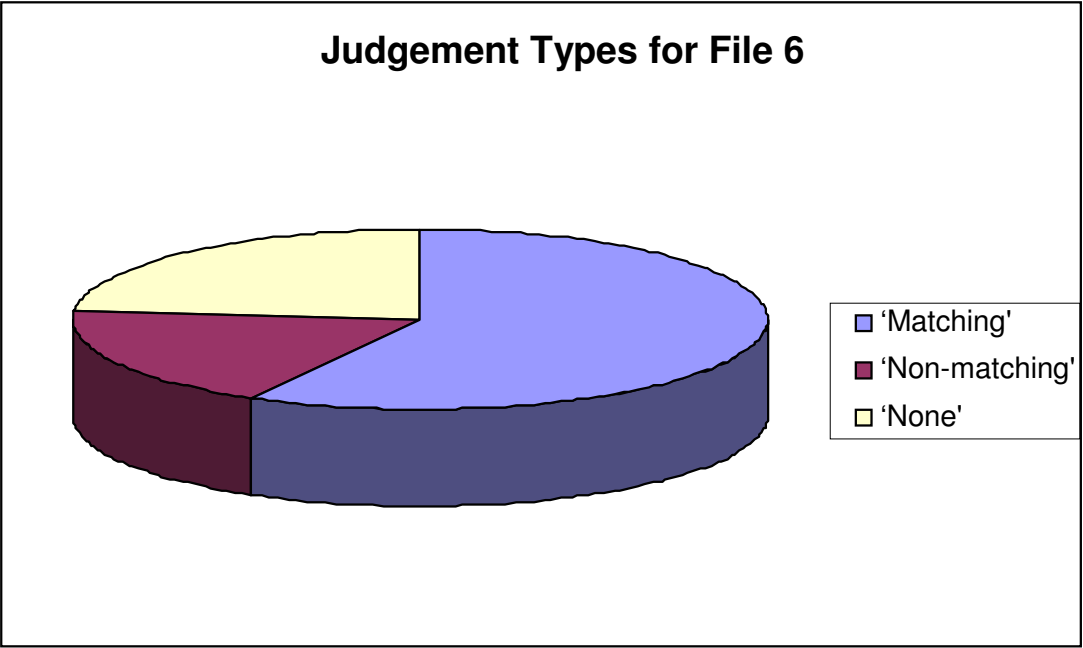
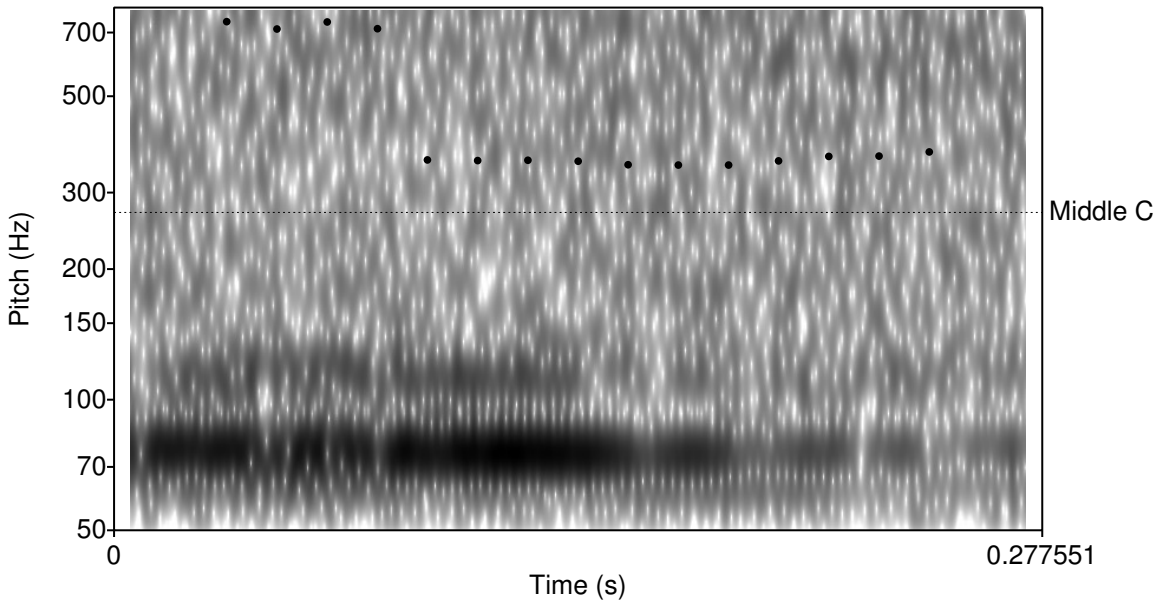
File 4: Is2, 23 week old, male infant, negative emotion (An 'eu' sound)



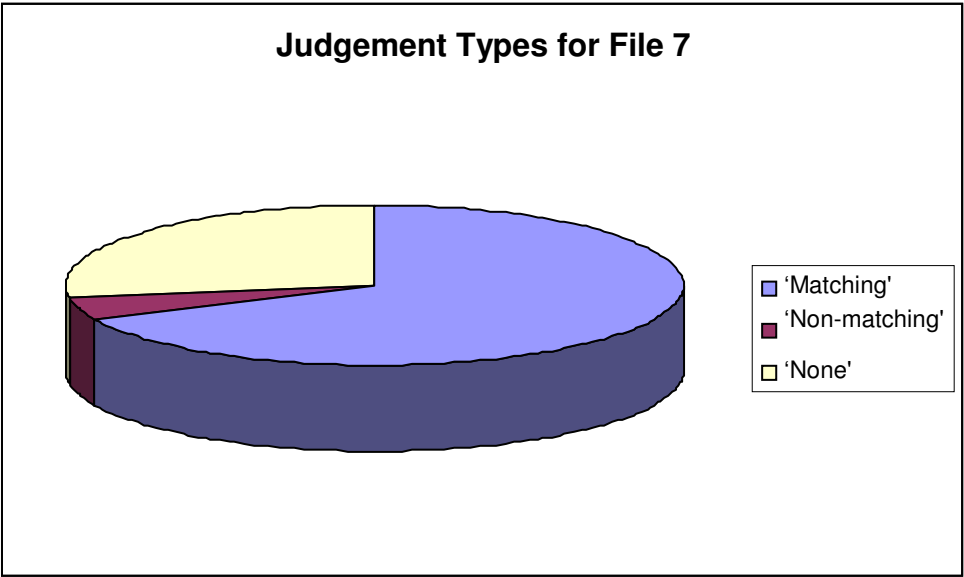
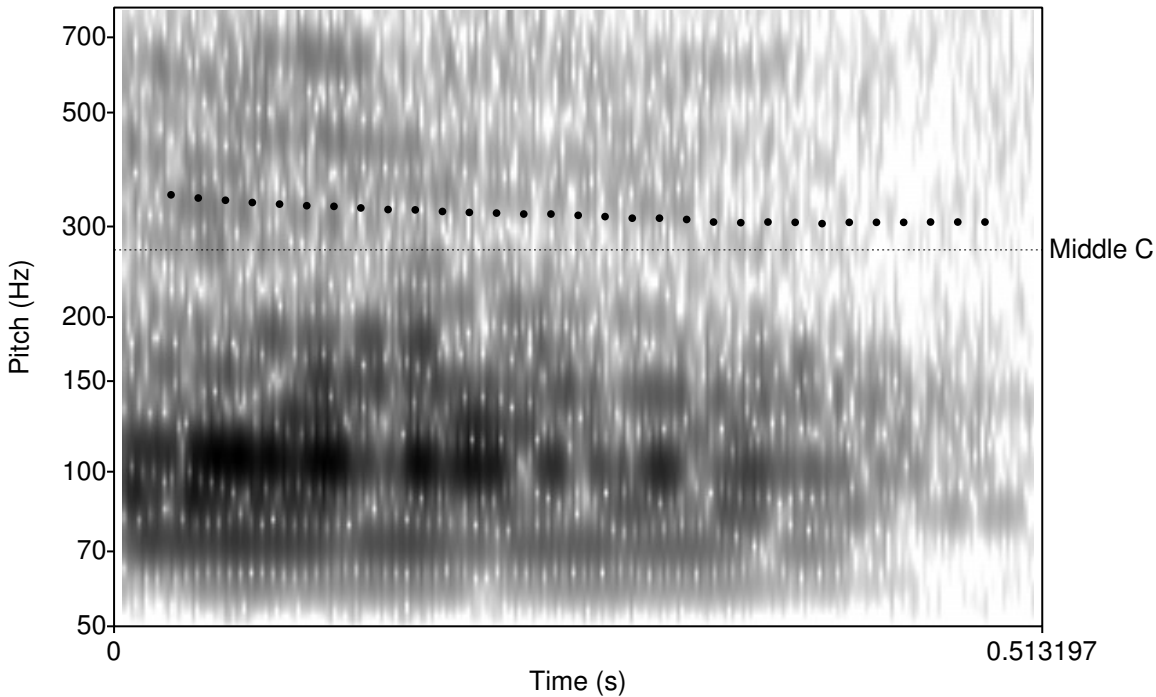
File 5: Jan 1, 28 week old, female infant, positive emotion (An long drawn out 'owue' sound)



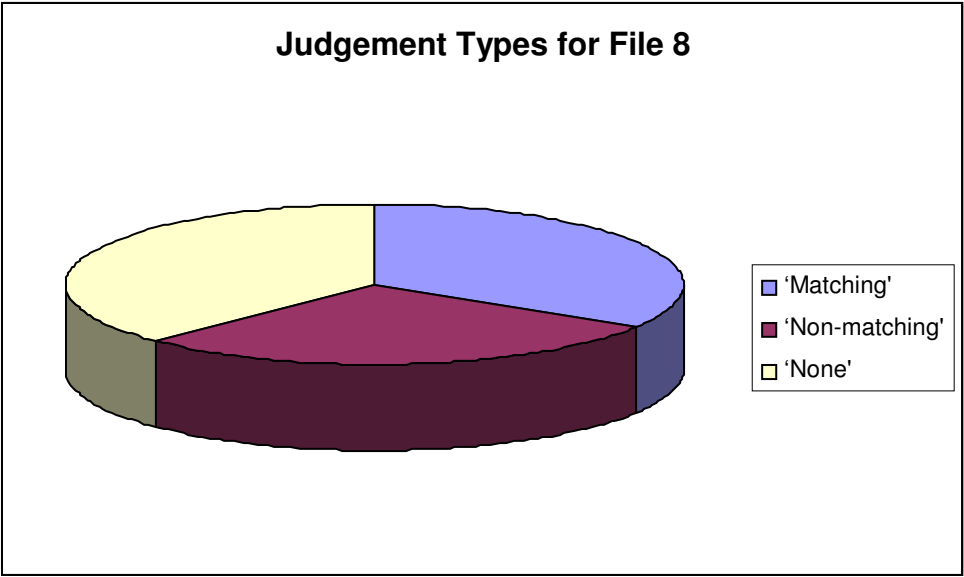
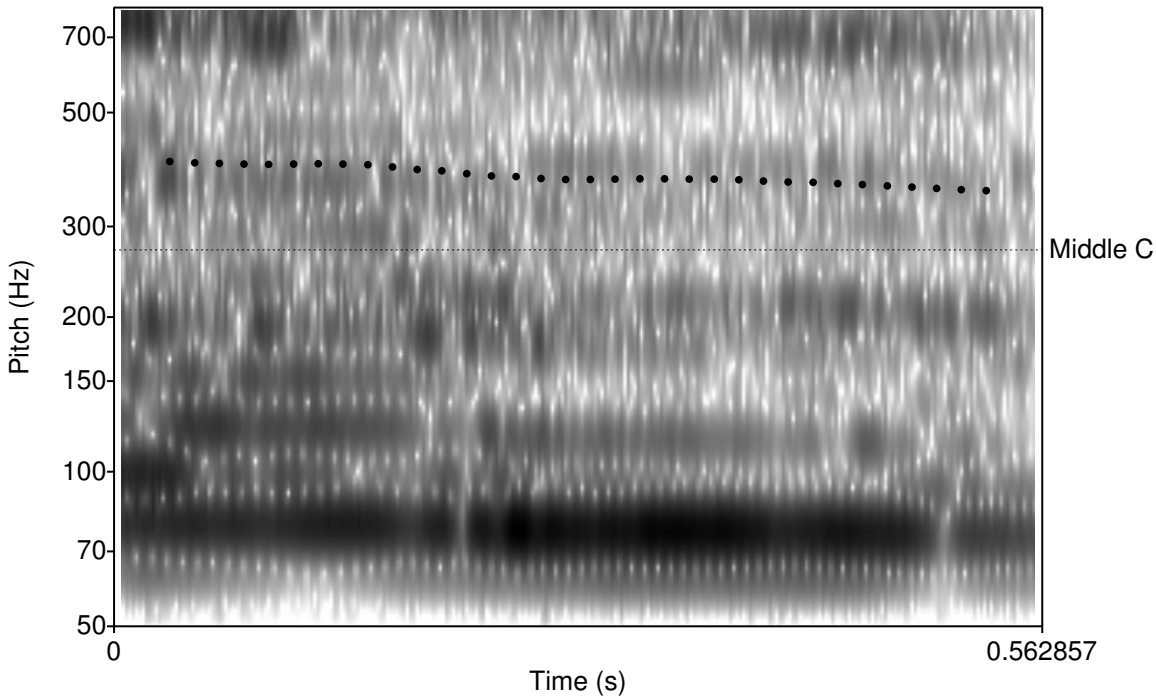
File 6: Jan 2, 28 week old, female infant, negative emotion (An ‘a-eu’ sound)



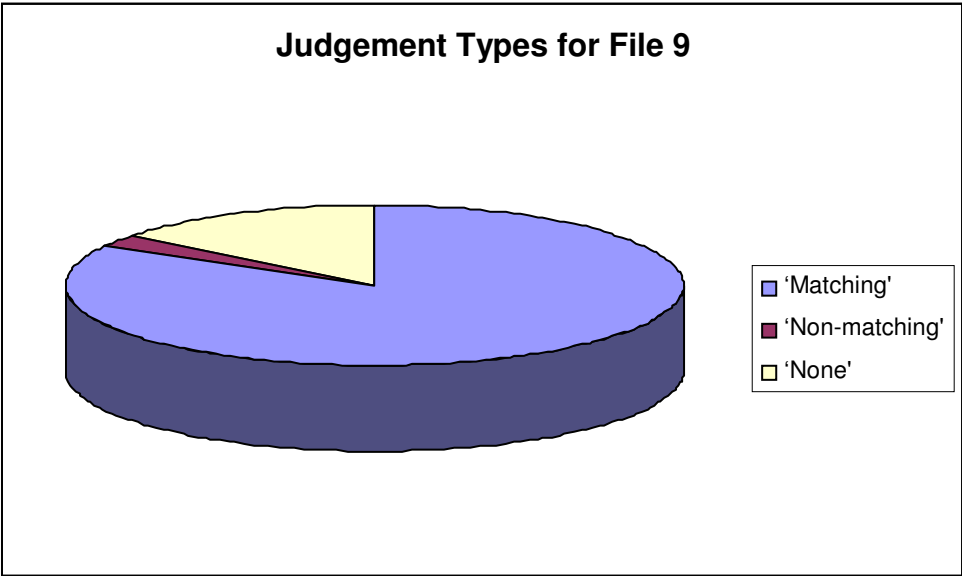
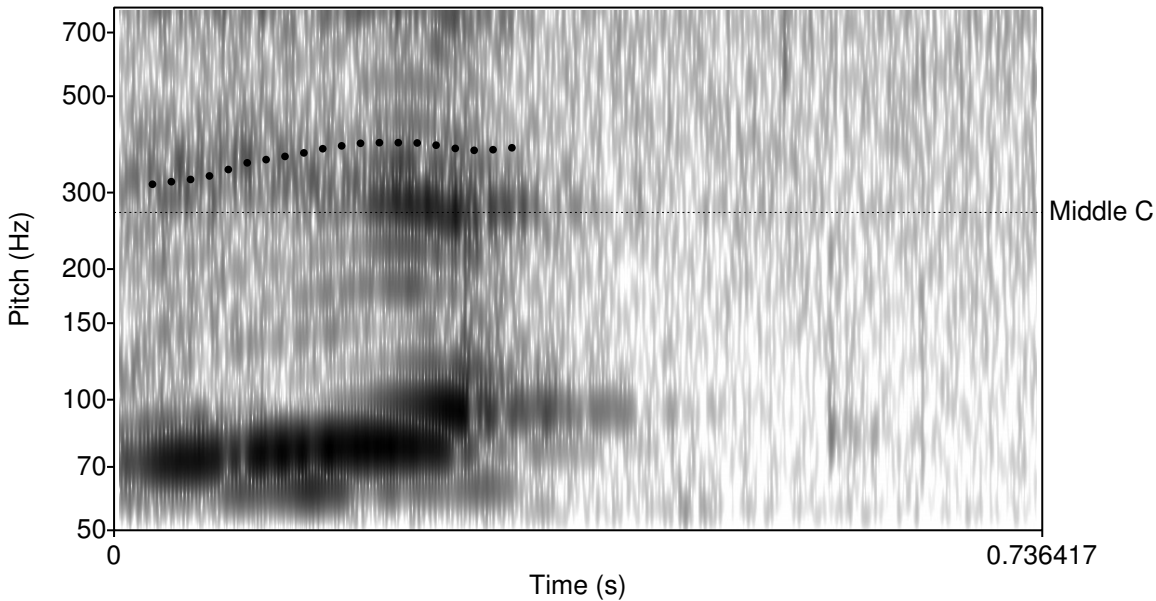
File 7: Ai 3, 36 week old, male infant, positive emotion (An ‘a’ sound)



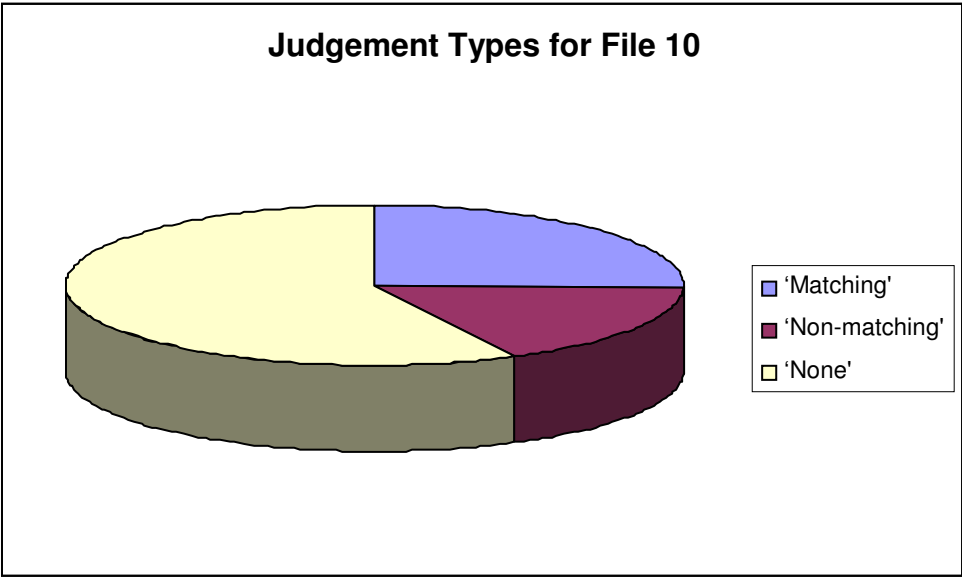
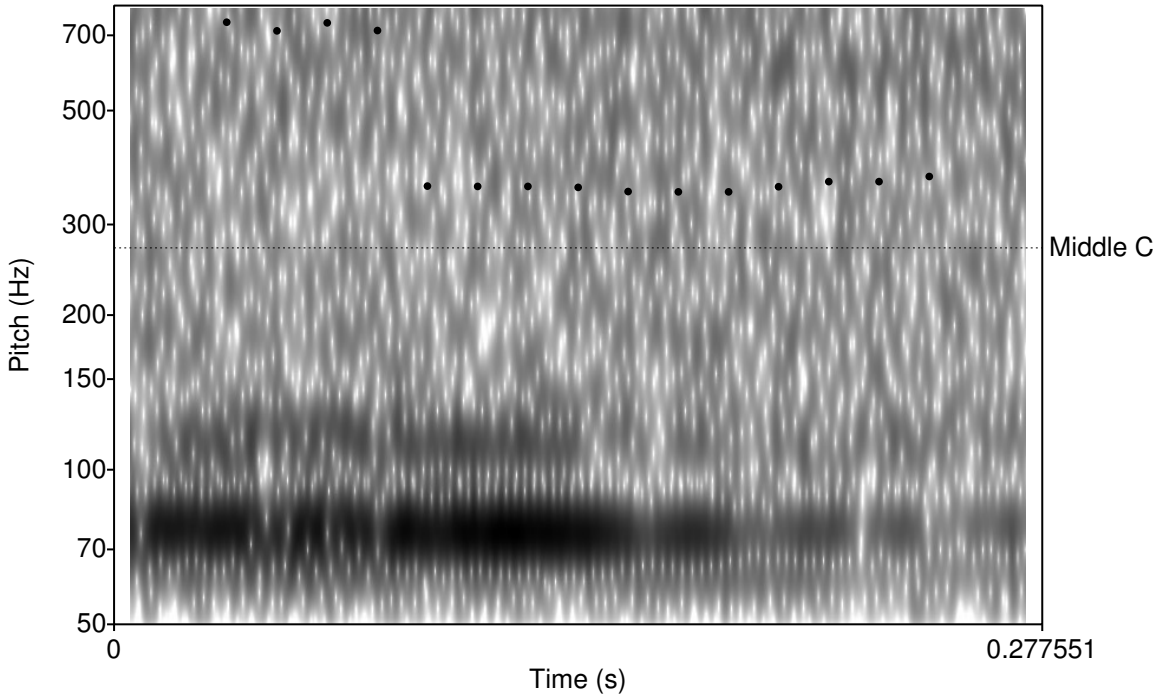
File 8: Ai 1, 36 week old, male infant, negative emotion ('ue' sound)



File 9: Li 1, 46 week old, female infant, positive emotion (A ‘ya’ sound)



File 10: Li 2, 46 week old, female infant, negative emotion (An ‘oh’ sound)



Appendix 2: Marwick and Trevarthen Category System for Affect

Interpersonal Affect

1. **Affectionate:** joyful, showing pleasure, in high spirits
2. **Empathetic:** echo or complete expression of affect, an immediate reflection of affect
3. **Sympathetic:** asking after feelings, commiserating
4. **Entranced:** riveted or enchanted by other action or production (including play) of other
5. **Amused/Enjoy:** enjoying an action or production of other
6. **Supportive:** passively following interest or action of other with positive affect and interest
7. **Approving:** congratulating on a display
8. **Appreciative:** recognition of a service
9. **Surprised:** astonished by, wonder at other
10. **Puzzled:** perplexed or confounded by other
11. **Bored:** wearied or fatigued by other
12. **Exasperated:** slightly irritated, upset or frustrated by other
13. **Disapproving:** of low opinion or disparaging of other
14. **Disgruntled:** disappointed, out of humour at other, dejected
15. **Depressed:** withdrawal from interpersonal communication
16. **Rejecting:** personal criticism, a break in the interpersonal relationship
17. **Disgusted:** at action or production from other
18. **Afraid:** suspicious or dismayed by other, wary
19. **Protesting:** angry, crying and vocalising, threshing about, struggling
20. **Aggressive:** kicking, biting, hitting or threatening

- 21. **Intractable:** resisting interpersonal engagement and persisting in self-directed action
- 22. **Overwhelming:** aggressive mocking, overfeedback
- 23. **Distress:** severely upset and sad

References

- Ainsworth, M.D.A. and Bell, S.M. (1970) Attachment, exploration, and separation: Illustrated by the behaviour of one-year-olds in a strange situation. Child Development. 41, 49-67.
- Andersen, P.A. and Guerrero, L.K. (1998) Principles of communication and emotion in social interaction. In Peter A. Anderson and Laura K. Guerrero (Eds), Handbook of Communication and Emotion: Research, Theory, Applications and Contexts p.49-89. New York, Academic Press.
- Bates, E. and Dick, F. (2002) Language, gesture and the developing brain. In: Developmental Psychobiology. 40, 393-310.
- Bateson, M.C. (1971) The Interpersonal Context of Infant Vocalization. Quarterly Progress Report of the Research Laboratory of Electronics. 100, 170-176.
- Beckman, M. (1996) The parsing of prosody. Language and Cognitive Processes. 11, 1-17.
- Beebe, B., Gerstman, L., Carson, B., Dolins, M., Zigman, A., Rosenweig, H., Faughey, K. and Korman, M. (1982) Rhythmic communication in the mother-infant dyad. In M. Davis (Ed), Interactional Rhythms: Periodicity in Communicative Behaviour p. 79-100. New York, Human Sciences Press.
- Beebe, B. and Gerstman, L. (1984) A method of defining 'packages of maternal stimulation and their functional significance for the infant with mother and stranger. International Journal of Behavioural Development. 7, 423-440.
- Beebe, B. and Lachmann, F.M. (2002) Infant Research and Adult Treatment:

Co-Constructing Interactions. London, Academic Press.

Bekoff, M. and Fox, M. W. (1972). Postnatal neural ontogeny: Environment-dependent and/or environment-expectant? Developmental Psychobiology, 5 (4), 323-341.

Bloom, K. and Masataka, N. (1996) Japanese and Canadian impressions of vocalising infants. International Journal of Behavioural Development. 19 (1), 89-99.

Boersma, P., Escurado, P. and Hayes, R. (2003) Learning Abstract Phonological from auditory phonetic categories: An integrated model for the acquisition of language-specific sound categories. Proceedings of the 15th International Congress of Phonetic Sciences. Barcelona. p.1013-1016.

Bornstein, M.H., Tamislemonda, C.S., Tal, J., Ludemann, P., Toda, S., Rahn, C.W., Pecheux, M.G., Azuma, H. and Vardi, D. (1992) Maternal responsiveness to infants in 3 societies – The United States, France and Japan. Child Development 63 (4): 808-821.

Bower, B. (1998) Getting a feel for emotions: Emotional development attracts cross-cultural explorations – science and emotions. In: Science News. 154 (12), 190-191.

Bowlby, J. (1969) Attachment and Loss: Volume One. Attachment. New York, Basic Books.

Bowlby, J. (1981) Attachment and Loss. Volume Three: Loss, Sadness, and Depression. New York, Basic Books.

Briggs, J.L. (1970) Never In Anger. Portrait of an Eskimo Family. Cambridge, Havard

University Press.

Bruner, J. S. and Sherwood, V. (1975) Early rule structure: the case of peekaboo.

In J. S. Bruner, A. Jolly and K. Sylva (Eds), Play: Its Role in Evolution and Development p.277-285. Harmondsworth: Penguin Books.

Bruner, J. (1983) Child's Talk. Oxford, Oxford University Press.

Bruner, J. (1990) Acts of Meaning. Press, Cambridge, MA, Harvard University.

Bruner, J. (1996) The Culture of Education. Cambridge, MA. Harvard University Press.

Buck, R. (1984) The Communication of Emotion. New York, The Guilford Press.

Byers, P. (1975) Biological rhythms as information channels in interpersonal communication behaviour. In P. Klopfer and G. Bateson (Eds), Perspectives in Ethology p.135-164. New York, Plenum.

Camras, L.A., DePaul, U., Meng, Z., Ujiie, T., Dharmusi, S., Miyake, K., Oster, H., Wang, L., Cruz, J., Murdoch, A. and Campos, J. (2002) Observing emotion in infants: Facial expression, body behaviour and rater judgements of responses to an expectancy-violating event. Emotion. 2 (2), 179-193.

Carvajal, F. and Iglesias, J. (2006) Judgements of facial and vocal signs of emotions in infants with Down's syndrome. Developmental Psychobiology. 48 (8), 644-652.

- Cheour, M., Alho, K., Sainio, K., Reinikainen, K., Renlund, M., Aaltonene, O., Eerola, O. and Naatanen, R. (1997) The Mismatch negativity to changes in speech sounds at the age of three months. Developmental Neuropsychology. 13 (2), 167-174.
- Chomsky, N. (1980) Rules and Representations. Oxford, Blackwell.
- Chomsky, N. (1996) On Nature and Language. Cambridge, Cambridge University Press.
- Clarkson, M. G., Martin, R. and Miciek, S. G. (1996) Infants' perception of pitch: Number of harmonics. In Infant Behaviour & Development. 19, 191-197.
- Cohen, A.J., Thorpe, L.A. and Trehub, S.E. (1987) Infants' perception of musical relations in short transposed tone sequences. Canadian Journal of Psychology. 41 (1), 33-47.
- Costa, Jr. P.T. and McCrae, R.R. (1992) The five-factor model of personality and its relevance to personality disorders. Journal of Personality Disorders. 6 (4), 343-359.
- Crown, C., Feldstein, S., Jasnow, M.D., Beebe, B. and Jaffe, J. (2002) The cross-modal coordination of interpersonal timing: Six-week-old infants' gaze with adults' vocal behavior. Journal of Psycholinguistic Research. 31 (1), 1-23.
- Crystal, D. (1997) The Cambridge Encyclopedia of Language (Second Edition). Cambridge, Cambridge University Press.
- Csibra, G. (2003) Teleological and referential understanding of action in

young infancy. Philosophical Transactions: Biological Sciences. 358 (1431), 447-458.

Darwin, C. (1872) The Expression of the Emotions in Man and Animals. London , John Murray.

Davis, B. L. and Lindblom, B. (2001) Phonetic variability in baby talk and development of vowel categories. In F. Lacerda, C. von Hofsten & M. Heimann (Eds), Emerging Cognitive Abilities in Early Infancy p.135-171. Mahwah NJ: Erlbaum.

Deliège, I., & Mélen, M. (1997) Cue abstraction in the representation of musical form. In I. Deliège & J. Sloboda (Eds), Perception and Cognition of Music p.387-412. Hove, Psychology Press.

DeCasper, A.J. and Fifer, W.P. (1980) Of human bonding: Newborns prefer their mother's voices. Science. 208, 1174-1176.

Dehaene-Lambertz, G., Hertz-Pannier, L. and Dubois, J. (2006) Nature and nurture in language acquisition: Anatomical and functional brain-imaging studies in infants. Trends in Neurosciences. 29 (7), 267-373.

Dewey, J. (2001) The theory of emotion. (1) Emotional Attitudes. Psychological Review 1, 553-569.

Diener, M.L. and Lucas, R.E. (2004) Adults desires for children's emotions across 48 countries: Associations with individual and national characteristics. Journal of Cross-Cultural Psychology. 35 (5), 525-547.

Dillard, J.P. (1998) Communication, biology and social relationships. In Peter A.

- Anderson and Laura K. Guerrero Eds), Handbook of Communication and Emotion: Research, Theory, Applications and Contexts p.xvii-xxxii. San Diego, Academic Press.
- Donald, M. (1999) Preconditions for the evolution of protolanguages. In M.C. Corballis and Lea, S.E.G Lea Eds), The Descent of Mind. Psychological Perspectives on Hominid Evolution p.355-365. Oxford, Oxford University Press.
- Donald, M. (2001) A Mind so Rare: The Evolution of Human Consciousness. New York, W.W. Norton and Company.
- Donaldson, M. (1992). Human Minds: An Exploration. London, Allen Lane/Penguin Books.
- Donohue, R.L. and Berg, W. K. (1991) Infant heart-rate responses to temporally predictable and unpredictable events. Developmental Psychology. 27 (1), 59-66.
- Draghi-Lorenz, R., Reddy, V. and Costall, A. (2001) Rethinking the development of “nonbasic” emotions: A critical review of existing theories. Developmental Review. 21 (3), 236-304.
- Dunn, J. (1999) Making sense of the social world: Mindreading, emotion and relationships. In P. D. Zelazo, J. Wilde Astington and D.R. Olson (Eds), Developing Theories of Intention: Social Understanding and Self-Control p229-242. London, Lawrence Erlbaum Associates.
- Durand, K., Lecuyer, R. and Frichtel, M. (2003) Representation of the third dimension: The use of perspective cues by 3- and 4-month-old infants. Infant Behaviour and Development. 26 (2), 151-166.
- Ejiri, K. and Masataka, N. (1999) Synchronization between preverbal vocal

- behaviour and motor action in early infancy II: An acoustical examination of the functional significance of the synchronization. Japanese Journal of Psychology. 69 (6), 433- 440.
- Ejiri, K and Masataka, N. (2001) Co-occurrence of preverbal vocal behaviour and motor action in early infancy. Developmental Science. 4 (1), 40-48.
- Ekman, P. (1972) Emotion in the human face : guidelines for research and an integration of findings . New York, Pergamon Press.
- Emde, R.N., Izard, C., Huebner, R., Sorce, J.F. and Klinnert, M. (1985) Adult judgements of infant emotions: Replication studies within and across laboratories. Infant Behaviour and Development. 8 (1), 79-88.
- Feldman, R. (2006) From biological rhythms to social rhythms: Physiological precursors of mother-infant synchrony. Developmental Psychology. 42 (1), 175-188.
- Feldstein, S., Jaffe, J., Beebe, B. and Crown, C. (1993) Coordinated interpersonal timing in adult-infant vocal interactions: A cross-site replication. Infant Behaviour and Development. 16, 455-470.
- Ferland, M.B. and Mendelson, M.J. (1989) Infants' categorization of melodic contour. Infant Behaviour and Development. 12 (3), 341-355.
- Fernald, A. & Morikawa, H. (1993) Common themes and cultural variations in Japanese and American mothers' speech to infants. Child Development. 64, 637-656.
- Fernald, A. (1989) Intonation and communicative intent in mother's speech to

- infants: Is the melody the message? Child Development 60, 1497-1510.
- Fernald, A. and Mazzie, C. (1991) Acoustic determinants of infant preference for motherese speech. Infant Behaviour and Development. 10 (3), 279-293.
- Fernald, A. (1993) Approval and disapproval: Infant responsiveness to vocal affect in familiar and unfamiliar languages. Child Development. 64, 657-674.
- Fifer, W.P. and Moon, C.M. (1995) The effects of fetal experience with sound. In J.P. Lecanuet, W.P. Fifer, N.A. Krasnegor and W.P. Smotherman (Eds), Fetal Development: A Psychobiological Perspective p.351-368. Hove, UK , Lawrence Erlbaum Associates.
- Fisher, C. and Tokura, H. (1996) Acoustic cues to grammatical structure in infant-directed speech: Cross-linguistic evidence. Child Development. 67 (6), 3192-3218.
- Fogel, A., Messinger, D.S., Dickson, K.L. and Hsu, H.C. (1999) Posture and Gaze in Early Mother-Infant Communication: Synchronization of Developmental Trajectories. Developmental Science. 2 (3), 325-332.
- Fogel, A., Nelson-Goens, G.C., Hsu, H.C. and Shapiro, A.F. (2000) Do different infant smiles reflect different positive emotion? Social Development . 9 (4), 497-520.
- Friedman, W.J. (2002) Arrows of time in infancy: The representation of temporal-causal invariances. Cognitive Psychology. 44 (3), 252-296.

- Friederici, A.D. and Wessels, J. M. I. (1993) Phonotactic knowledge of word boundaries and its use in infant speech-perception. Perception and Psychophysics, 54 (3), 287-295.
- Gergely, G. and Watson, J.S. (1999) Early socio-emotional development: Contingency perception and the social-biofeedback model. In P. Rochat (Ed), Early social cognition: Understanding others in the first months of life p. 101-136. Hove, Lawrence Erlbaum Associates.
- Geertz, C. (2000) Available Light: Anthropological Reflections on Philosophical Topics. New Jersey, Princeton University Press.
- Gibson, J.J. (1979) The Ecological Approach to Visual Perception. Boston, Houghton Mifflin.
- Goldfield, E.C. (2000) Exploration of vocal tract properties during serial production of vowels by full term and preterm infants. Infant Behaviour and Development 23 (3-4): 421-439.
- Gorski, P.A. (2001) Contemporary paediatric practise: In support of infant mental health (imaging and imagining). Infant Mental Health Journal 22 (1-2), 188-200.
- Gratier, M. (1999) Expressions of belonging: The effect of acculturation on the rhythm and harmony of mother-infant vocal interaction. Musicae Scientiae, Special Issue, 1999-2000, 93-122.
- Gratier, M. (2003) Protoconversation and the Roots of Belonging. Centre for Culture, Brain and Development. Los Angeles.
- Gratier, M. (2006) Finding and growing roots: The musical dynamics of

belonging. Submitted for publication in collected papers in honour of Mikis Theodorakis.

Green, J.A., Gustafson, G.E., Irwin, J.R., Kalinowski, L.L. and Wood, R.M.

(1995) Infant crying: Acoustics, perception and communication. Early Development and Parenting. 4 (4), 161-175.

Guedeney, A. and Fermanian, J. (2001) A validity and reliability study of assessment and screening for sustained withdrawal reaction in infancy: The alarm distress baby scale. Infant Mental Health Journal. 22(5), 559-575.

Hauser, M.D., Chomsky, N. and Fitch, W.T. (2002) The faculty of language: What is it, who has it, and how did it evolve? Science. 298 (5598), 1569-1579.

Hauser, M.D. and McDermott, J. (2003) The evolution of the music faculty: A comparative perspective. Nature Neuroscience. 6 (7), 663-668.

Heimann, M. (2001) Neonatal imitation – A fuzzy phenomenon? In F. Lacerda, C. von Hofsten and Heimann, M. (Eds), Emerging Cognitive Abilities in Early Infancy p.231-246. Hove, Lawrence Erlbaum.

Henderson, H.A., Fox, N.A. and Rubin, K.H. (2001) Temperamental contributions to social behaviour: The moderating roles of frontal EEG asymmetry and gender. Journal of the American Academy of Child and Adolescent Psychiatry 40 (1), 68-74.

Hobson, P. (2004) The Cradle of Thought: Exploring the Origins of Thinking. London, Pan Books.

Hobson, P. (2007) Communicative depth: Soundings from developmental psychology. Infant Behaviour and Development. 30(2), 267-277.

Horner, T.M. and Chetnik, L. (1986) Conversation attentiveness and following in 12-

- and 18-week old infants. Infant Behaviour and Development. 9, (2), 203-213.
- Houstén, D.M., Jusczyk, P.W., Kuijpers, C., Coolen, R. and Cutler, A. (2000) Cross-language word segmentation by 9-month-olds. Psychonomic Bulletin and Review. 7 (3), 504-509.
- Houstén, D.M. and Jusczyk, P.W. (2000) The role of talker-specific information in word segmentation by infants. Journal of Experimental Psychology – Human Perception and Performance. 26 (5), 1570-1582.
- Hsu, H.C., Fogel, A. and Cooper, R.B. (2000) Infant vocal development during the first 6 months: Speech quality and melodic complexity. Infant and Child Development. 9 (1), 1-16.
- Hsu, H.C. and Fogel, A. (2001) Infant vocal development in a dynamic mother-infant communication system. Infancy. 2 (1), 87-109.
- Hsu, H.C., Fogel, A. and Messinger, D. (2001) Infant non-distress vocalization during mother-infant face-to-face interaction: Factors associated with quantitative and qualitative differences. Infant Behaviour and Development. 24 (1), 107-128.
- Ikeda, Y. and Masataka, N. (1999) A variable that may affect individual differences in the child-directed speech of Japanese women. Japanese Psychological Research. 41 (4), 203-208.
- Issartel, J., Marin, L., Gayraud, J., Hauw, D., Mottet, D., and Cadopi, M. (2004) Nature of motor interaction in a specific dance task: The improvisation. 8th European Workshop on Ecological Psychology. Verona, Italy. June.

- Iverson, J.M., Hall, A.J., Nickel, L. and Wozniak, R.H. (2006) The relationship between reduplicated babble onset and laterality biases in infant rhythmic arm movements. Brain and Language. 101 (3), 198-207.
- Izard, C.E., Huebner, R.R., Risser, D. and Dougherty, L. (1980) The Young Infant's Ability to Produce Discrete Emotion Expressions. In: Developmental Psychology. Vol 16 (2), p. 132-140.
- Jaffe, J., Beebe, B., Feldstein, S., Crown, C. and Jasnow, M. (2001) Rhythms of dialogue in infancy: Coordinated timing and social development. SRCD Monographs, 66, (2) Serial No. 264, 1-132.
- Jonsson, C., Clinton, D., Fahrman, M., Mazzaglia, G., Novak, S. and Sörhus, K. (2001) How do mothers signal shared feeling-states to their infants? An investigation of affect attunement and imitation during the first year of life. Scandinavian Journal of Psychology. 42 (4), 377-381.
- Jusczyk, P.W. and Luce, P.A. (1994) Infants' sensitivity to phonotactic patterns in the native language. Journal of Memory and Languages. 33 (5), 630-645.
- Jusczyk, P.W. (1997) The role of infant speech perception capacities in discovering the sound structure of the native language (invited paper). 133rd Meeting of Acoustical Society of America. (Report) Journal of the Acoustical Society of America. 101 (5), 3192.
- Jusczyk, P.W. (1998) Developing sensitivity to native language sound patterns (invited paper). 133rd Meeting of Acoustical Society of America. (Report) Journal of the Acoustical Society of America. 103 (5), 2931.

- Jusczyk, P.W. (1999) How infants begin to extract words from speech. Trends in Cognitive Sciences. 3 (9), 323-328.
- Jusczyk, P. W. (2001) In the beginning was the word. In F. Lacerda, C. von Hofsten and M. Heimann (Eds) Emerging Cognitive Abilities in Early Infancy p. 173-192. Mahwah NJ, Erlbaum.
- Kahana-Kalman, R. and Walker-Andrews, A.S. (2001) The role of person familiarity in young infants' perception of emotional expressions. Child Development. 72 (2), 352-369.
- Kaitz, M. and Maytal, H. (2005) Interactions between anxious mothers and their infants: An integration of theory and research findings. Infant Mental Health Journal. 26 (6), 570-597.
- Keller, H. and Schölmerich, A. (1987) Infant Vocalizations and Parental Reaction During the First 4 Months of Life. In: Developmental Psychology. 23 (1), 62-67.
- King-Cline, S. (1987) "Round and round the garden": An analysis of nursery rhyme song in mother-infant interaction." Unpublished Honours Thesis, Department of Psychology, University of Edinburgh.
- Kitamura, C. and Burnham, D. (1998) The infant's response to maternal vocal affect. In C. Rovee-Collier, L.P. Lipsitt and H. Hayne (Eds), Advances in Infancy Research Chapter 12, p. 221-236. Stamford, Ablex.
- Kohler, E., Keysers, C., Umiltà, M.A., Fogassi, L., Gallese, V. and Rizzolatti, G.

- (2002) Hearing sounds, understanding actions: Action representation in mirror neurons. Science, 297, 846-848.
- Kugiumutzakis, G. (1993) Intersubjective vocal imitation in early mother-infant interaction. In J. Nadel and L. Camioni (Eds), New Perspectives in Early Communicative Development p. 23 – 47. London, Routledge.
- Kugiumutzakis, G., Kokkinaki, T. Markodimitraki, M. and Vitalaki, E. (2005). Emotions in early mimesis. In J. Nadel and D. Muir (Eds), Emotional Development p. 161-182. Oxford, Oxford University Press.
- Kuhl, P. K. (1983) Perception of auditory equivalence classes for speech in early infancy. Infant Behaviour and Development. 6(2-3), 263-285.
- Kuhl, P. K. (1994) Learning and representation in speech and language. Current Opinion in Neurobiology. 4 (6), 812-822.
- Kuhl, P.K. and Meltzoff, A.N. (1996) Infant vocalisations in response to speech: Vocal imitation and developmental change. Journal of Acoustical Society of America. 100 (4), 2425-2438.
- Kuhl, P.K., Andruski, J.E., Chistovich, L.A., Kozhevnikova, E.V., Ryskina, V.L., Stolyarova, E.L., Sundberg, U. and Lacerda, F. (1997) Cross-language analysis of phonetic units in language addressed to infants. Science. 277 (5326), 684-686.
- Kuhl, P. K. (1998) Language, culture and intersubjectivity: The creation of shared perception. In S. Bråten (Ed), Intersubjective Communication and Emotion in Early Ontogeny p.307-315. Cambridge, Cambridge University Press.

Kuhl, P. K. (2001) Speech, language, and developmental change. In F. Lacerda, C. von Hofsten & M. Heimann (Eds), Emerging Cognitive Abilities in Early Infancy p.111-133. Mahwah, NJ, Erlbaum.

Lacerda, F. & Sundberg, U. (2001) Auditory and articulatory biases influence the initial stages of the language acquisition process. In F. Lacerda, C. von Hofsten & M. Heimann (Eds), Emerging Cognitive Abilities in Early Infancy p.91-110. Mahwah, NJ, Erlbaum.

Ladd, D.R., Silverman, K.E., Tolkmitt, F., Bergmann, G. and Scherer, K.R. (1985) Evidence for the independent function of intonation contour type, voice quality, and FO range in signalling speaker affect. Journal of the Acoustical Society of America. 78 (2), 435-444.

Ladd, D.R. (1996) Intonational Phonology. Cambridge, Cambridge University Press.

Lecanuet, J.P., Fifer, W. P., Krasnegor, N. A. and Smotherman, W. P. (1996) Fetal Development: A Psychobiological Perspective. Hillsdale, NJ, Erlbaum.

Lee, D.N. (1998) Guiding Movement by Coupling Taus. Ecological Psychology. 10 (3-4), 221-250.

Legerstee, M. (1997) Contingency effects of people and objects on subsequent cognitive functioning in three-month-old infants. Social Development. 6 (3), 307-321.

Legerstee, M. (2001) Six-month-old infants rely on explanatory inference when relating

communication to people and manipulatory actions to inanimate objects: Reply to Gergely. Developmental Psychology. 37 (5), 583-586.

Leveille, E., Cossette, L., Blanchette, I. And Gaudreau, M. (2001) The socialisation of emotions during nursing: The role of contingent maternal facial expressions. International Journal of Psychology. 36 (4), 260-273.

Lewis, C. (1995) Educating Hearts and Minds. Cambridge, Cambridge University Press.

Locke, J. L. (1993) The Child's Path to Spoken Language. Cambridge MA and London, Harvard University Press.

Locke, J.L. (1994) The biological building blocks of spoken language. In J.A. Hogan and J. J. Bolhuis (Eds), Causal Mechanisms of Behavioural Development p. 300-324. New York, Wiley.

Loughlin, E. (1999) Shared dance: Dance therapy with mothers and infants in the hospital outpatient infant clinic. Dance Therapy Collections. 2, 37-42

Malatesta, C.Z. (1981) Infant emotion and the vocal affect lexicon. Motivation and Emotion. 5 (1), 1-23.

Mangelsdorf, S., McHale, J.L., Diener, M., Heim Goldstein, L. and Lehn, L. (2000) Infant attachment: Contributions of infant temperament and maternal characteristics. Infant Behaviour and Development. 23, 175-196.

Masataka, N. (1992) Early ontogeny of vocal behaviour of Japanese infants in

- response to maternal speech. Child Development. 63, 1177-1185.
- Masataka, N. (1992) Motherese in a signed language. Infant Behaviour and Development. 15, 453-460.
- Masataka, N. (1993) Relation between pitch contour of prelinguistic vocalisations and communicative functions in Japanese infants. Infant Behaviour and Development. 16 (3), 397-401.
- Masataka, N. (1995) The relation between index-finger extension and the acoustic quality of cooing in 3-month-old infants. Journal of Child Language. 22 (2), 247-257.
- Masataka, N. (1999) Preference for infant-directed singing in a 2-day-old hearing infants of deaf parents. Developmental Psychology. 35 (4), 1001-1005.
- Matthews, G. and Deary, I.J. (1998) Personality Traits. Cambridge, Cambridge University Press.
- Mattys, S.L. and Jusczyk, P.W (2001) Do infants segment words or recurring contiguous patterns? Journal of Experimental Psychology – Human Perception and Performance. 27 (3): 644-655.
- Malloch, S., Sharp, D., Campbell, D. M., Campbell, A.M. & Trevarthen, C. (1997) Measuring the human voice: Analysing pitch, timing, loudness and voice quality. Proceedings of the Institute of Acoustics, 19, 5 (495-500).
- Malloch, S. (1999). Mothers and infants and communicative musicality.

Musicae Scientiae, Special Issue, 1999-2000, p. 29-57.

Mehler, J. (2000) See <http://www.cog.brown.edu/~morgan/> (8th July. 2004).

Miall, D.S. and Dissanayake, E. (2002) The poetics of babytalk. Human Nature. 14 (4), 337-364.

Milligan, K., Atkinson, L., Trehub, S., Benoit, D. and Poulton, L. (2003) Maternal attachment and the communication of emotion through song. Infant Behaviour and Development. 26, 1-13.

Morgan, J.L. (1994) Converging measures of speech segmentation in preverbal infants. Infant Behaviour and Development. 17 (4), 389-403.

Morgan, J. and Saffran, J.R. (1995) Emerging integration of sequential and suprasegmental information in preverbal speech segmentation. Child Development. 66, 911-936.

Murray, L. (1980) The Sensitivities and Expressive Capacities of Young Infants in Communication with their Mothers. Doctoral Dissertation. University of Edinburgh.

Murray, L. (1992) The impact of postnatal depression on infant development. Journal of Child Psychology and Psychiatry 34, 1083-1101.

Murray, L., Kempton, C., Woolgar, M. and Hooper, R. (1993) Depressed mothers' speech to their infants and it's relation to infant gender and cognitive development. Journal of Child and Psychiatry 34 (7), 1083-1101.

- Murray, L. & Trevarthen, C. (1985) Emotional regulation of interactions between two-month-olds and their mothers. In Field, T. and Fox, N. (Eds), Social Perception in Infants p.177-197. Norwood, N.J., Ablex.
- Nadel, J., Carchon, I., Kervella, C., Marcelli, D., and Réserbat-Plantey, D. (1999). Expectancies for social contingency in 2-month-olds. Developmental Science, 2 (2), 164-173.
- Nadel, J. Carchob, I., Kervella, C., Marcelli, D. and Reserbat-Plantey, D. (1999) Expectancies for social contingency in 2-month-olds. Developmental Science. 2, 164-173.
- Nadel, J. (2004) When Imitation is Awakened in Children With Autism.
Keynote Lecture, BPS Developmental Conference, Leeds. September.
- Nagy, E. and Molnar, P. (1994) Homo imitans or homo provocans? Abstracts 7th IOP Scientific Meeting/International Journal of Psychophysiology. 18 (2),128.
- Nagy, E. and Molnar, P. (2004) Homo imitans or homo provocans? The phenomenon of neonatal initiation. Infant Behaviour and Development. 27, 57-63.
- Nagy, E. (2004) From imitation to conversation: The first dialogues with human neonates. Infant and Child Development. 15 (3), 223-232.
- Nagy, E., Kompagne, H., Orvos, H. and Pal, A. (2007) Gender-related differences in neonatal imitation. Infant and Child Development. 16 (3), 267-276.
- Nakano, S. (1997) Heart-to-heart (inter-jo-) resonance: A concept of

- intersubjectivity in Japanese everyday life. Annual Report of Research and Clinical Centre of Child Development. 19, 1-14.
- Nazzi, T., Jusczyk, P.W. and Johnson, E.K. (2000) Language discrimination by English-learning 5-month-olds: Effects of rhythm and familiarity. Journal of Memory and Language. 43 (1), 1-19.
- Neisser, U. (1994) Self perception and self-knowledge. Psyke and Logos 15, 392-407.
- Nelson, H., Pick S. and Pick, A.D. (1999) Biobehavioural development, perception and action. Infants' perception of dynamic affective expressions: Do infants distinguish specific expressions? Child Development. 70(6), 1275.
- Oller, D.K. (1986) Metaphonology and infant vocalisations. In B. Lindblom and R. Zetterstrom (Eds), Precursors of Early Speech p.21-35. Basingstoke, Macmillan.
- Ochsner, K. and Barrett, L. (2001) A multiprocess perspective on the neuroscience of emotion. In T. Mayne and G. Bonanno (Eds), Emotions: Current Issues and Future Directions. Emotions and Social Behaviour p.38-81. New York, The Guildford Press.
- Palmer, C., Jungers, M.K. and Jusczyk, P.W. (2001) Episodic memory for musical prosody. Journal of Memory and Language. 45 (4), 526-545.
- Panksepp, J. (2001) The long term psychobiological consequences of infant emotions: Prescriptions for the twenty-first century. Infant Mental Health Journal 22 (1-2), 132-173.
- Panksepp, J. (2001) At the interface of the affective, behavioural and cognitive

neurosciences: Decoding the emotional feelings of the brain. Brain and Cognition. 52 (1), 4-14.

Panksepp, J. and Bernatzky, G. (2002) Emotional sounds and the brain: the neuro-affective foundations of musical appreciation. Behavioural Processes. 60,133-135.

Papaeliou, C.F., Midadakakis, G. and Cavouras, D. (2002) Acoustic patterns of infant vocalizations expressing emotions and communicative functions. Journal of Speech, Language and Hearing Research. 45 (2), 311-317.

Papaeliou, C.F. and Trevarthen, C. (2006) Prelinguistic pitch patterns expressing 'communication' and 'apprehension'. Journal of Child Language. 33 (1), 163-178.

Papoušek, M. and Papoušek, H. (1981). Musical elements in the infant's vocalization: Their significance for communication, cognition, and creativity. In L. P. Lipsitt and C. K. Rovee-Collier (Eds), Advances in Infancy Research, 1, p.163-224. Norwood, NJ, Ablex.

Papoušek, H. and Papoušek, M. (1986) Structure and dynamics of human communication at the beginning of life. European Archives of Psychiatric Neuological Science. 236, 21-25.

Papousek, M., Bornstein, M., Nuzzo, C., Papousek, H. and Symmes, D. (1990) Infant responses to prototypical melodic contours in parental speech. Infant Behaviour and Development. 13 (4), 539-545.

Papoušek, M. and Papoušek, H. (1991) The meanings of melodies in motherese

in tone and stress languages. Infant Behaviour and Development. 14, 415-440.

Papoušek, M. (1992) Early Ontogeny of Vocal Communication in Parent-Infant Interactions. In H. Papoušek, U. Jurgens and M. Papoušek (Eds), Noverbal Vocal Communication. Comparative and Developmental Approaches. Studies in Emotion and Social Interaction p230-261. Cambridge, Cambridge University Press.

Papoušek, H. and Bornstein, M.H. (1992) Didactic interactions: Intuitive parental support of vocal and verbal development in human infants. In H.Papousek, U. Jurgens and M. Papousek (Eds), Noverbal Vocal Communication. Comparative and Developmental Approaches. Studies in Emotion and Social Interaction p.209-229. Cambridge, Cambridge University Press.

Papoušek, M. (1996). Intuitive parenting: A hidden source of musical stimulation in infancy. In Delige and Sloboda (Eds), Musical beginnings : origins and development of musical competence p.88-112. Oxford, Oxford University Press.

Papousek, M. (2007) Communication in early infancy: An arena of intersubjective learning. Infant Behaviour and Development. 30, 258-266.

Parise, E., Cleveland, A., Costabile, A. and Striano, T. (2007) Influence of vocal cues on learning about objects in joint attention contexts. Infant Behaviour and Development. 30 (2), 380-384.

Pelchat, D., Bisson, P., Bois, C. and Saucier, J-F. (2003) The effects of early relational antecedents and other factors on the parental sensitivity of mothers and fathers. Infant and Child Development. 12 (1), 27-51.

Pesonen, A-K., Räikkönen, K., Strandberg, T., Letlikangas-Järvinen, L and Järvenpää,

- A-L. (2004) Insecure adult attachment style and depressive symptoms: Implications for parental perceptions of infant temperament. Infant Mental Health Journal. 25 (2), 99-116.
- Pierrehumbert, J. B. & Beckman, M. E. (1988) Japanese Tone Structure. Cambridge, MA, MIT Press.
- Planalp, S. (1998) Communicating Emotion in Everyday Life: Cues, Channels and Processes. In Peter A. Anderson and Laura K. Guerrero (Eds), Handbook of Communication and Emotion: Research, Theory, Applications and Contexts p.30-45. New York, Academic Press.
- Polka, L and Bohn, O.S. (1996) Cross-language comparison of vowel perception in English-learning and German-learning infants. Journal of The Acoustic Society of America. 100 (1), 577-592.
- Pollack, S.D., Cicchetti, D., Hornung, K. and Reed, A. (2000) Recognising emotion in faces: Developmental effects of child abuse and neglect. Developmental Psychology 36 (5), 679-688.
- Port, R., Dalby, J., & O'Dell, M. (1987) Evidence for mora timing in Japanese. Journal of the Acoustical Society of America. 81, 1574-1585.
- Powers, N. (2001) Intrinsic musicality: Rhythm and prosody in infant-directed voices. In: Annual Report, 1999-2000, Research and Clinical Center for Child Development, Faculty of Education, Hokkaido University, Japan. 23, 1-19.
- Pulvermüller, F. (1999) Words in the brain's language. Behavioural and Brain Sciences. 22, 253-336.

- Quinn, P.C. (2002) Category representation in young infants. Current Directions in Psychological Science. 11 (2), 66-70.
- Quinn, P.C. and Schyns, P.G. (2003) What goes up may come down: Perceptual process and knowledge access in the organisation of complex visual patterns by young infants. Cognitive Science. 27 (6), 923-935.
- Ramus, F., Nespor, M. and Mehler, J. (1999) Correlates of linguistic rhythm in the speech signal. Cognition. 73, 265-292.
- Ramus, F., Hauser, M.D., Miller, D., Morris, D. and Mehler, J. (2000) Language discrimination by human newborns and cotton-top tamarins. Science. 288, 349-351.
- Reddy, V., Hay, D., Murray, L., & Trevarthen, C. (1997). Communication in infancy: Mutual regulation of affect and attention. In G. Bremner, A. Slater, & G. Butterworth (Eds), Infant Development Recent Advances p. 247-275. Hove, Psychology Press.
- Reddy, V. (2002) Rhythms in Action and Emotion Workshop Introduction. (Oral presentation). University of Portsmouth, Portsmouth, UK. June.
- Reddy, V. and Trevarthen, C. (2004) What we learn about babies from engaging with their emotions. Zero to Three, January 2004. 24 (3), 9-15.
- Reznick, J.S. (1999) Influences on maternal attributions of infant intentionality. In P.D.

Zelazo, J. Wilde Ashington and D.R. Olson (Eds), Developing Theories of Intention: Social Understanding and Self-Control p.243-267. London, Lawrence Erlbaum Associates Publishers.

Rizzolatti, G. and Arbib, M.A. (1998) Language within our grasp. Trends in Neurosciences. 21, 188 - 194.

Rizzolatti, R., Fogassi, L. & Gallese, V. (2001) Neurophysiological mechanisms underlying the understanding and imitation of action. Nature Reviews Neuroscience. 2, 661 -670.

Rizzolatti, G. and Craighero, L. (2004) The mirror-neuron system. Annual Review of Neurosciences. 27, 169 – 192.

Robb, L. (1999) Emotional musicality in mother-infant vocal affect, and an acoustic study of postnatal depression. Musicae Scientiae. Special Issue 1999-2000, 123-154.

Rochat, P. and Striano, T. (1999) Social-cognitive development in the first year. In P. Rochat (Ed), Early Social Cognition: Understanding Others in the First Months of Life p. 3-34. Mahwah, New Jersey, Erlbaum.

Rochat, P., Querido, J.G. and Striano, T. (1999) Emerging sensitivity to the timing and structure of protoconversation in early infancy. Developmental Psychology. 35 (4), 950-957.

Rochat, P. and Striano, T. (2000) Perceived self in infancy. Infant Behaviour and Development. 23, 513-530.

- Rochat, P. (2001a) The Infant's World. Cambridge, M.A, Harvard University Press.
- Roe, K.V. (2001) Relationship between male infants' vocal responses to mother and stranger at three months and self-reported academic attainment and adjustment measures in adulthood. Psychological Reports 89 (2), 255-258.
- Ross, D., Choi, J. and Purves, D. (2007) Musical intervals in speech. Proceedings of National Academy of Sciences of United States of America. 104 (23), 9852-9857.
- Sansavini, A. (1997) Neonatal perception of the rhythmical structure of speech: The role of stress patterns. Early Development and Parenting. 6 (1), 3-13.
- Scherer, K.R. (1992) Vocal affect expression as symptom, symbol and appeal. In H. Papoušek, U. Jürgens, M. Papoušek (Eds), Nonverbal Vocal Communication. p.43-60. Cambridge, Cambridge University Press.
- Schegloff, E.A. (1998) Reflections on studying prosody in talk-in-interaction. Language and Speech: Special Double Issue on Prosody and Conversation. 41 (3-4), 235-263.
- Schögler, B. (1999) Studying temporal co-ordination in jazz duets. Musicae Scientiae. Special Issue 1999-2000, 75-92.
- Schoppe-Sullivan, S.J., Diener, M.L., Mangelsdorf, S.C., Brown, G.L., McHale, J.L. and Frosch, C.A. (2006) Attachment and sensitivity in family context: The roles of parent and infant gender. Infant and Child Development. 15 (4), 367-385.

Schore, A. (2000) Attachment and the regulation of the right brain. Attachment and Human Development. 2 (1), 23-47.

Schore, A. (2002) Dysregulation of the right brain: a fundamental mechanism of traumatic attachment and the psychopathogenesis of posttraumatic stress disorder. Australian and New Zealand Journal of Psychiatry. 36 (1), 9-30.

Shibukawa, Y., Shintani, M., Kumai, T., Kato, M., Suzuki, T., Zhang, Z., Jiang, T., Shimono, M., Ishikawa, T. and Nakamura, Y. (2003) Jaw-Movements-Related Mirror Neuron System in Humans. Poster presented at 81st General Session of the International Association for Dental Research. Gothenborg, Sweden. June.

Shimura, Y. and Imaizumi, S. (1995) Emotional information in young infants' vocalisations. Proceedings of the International Congresses of Phonetic Sciences. 3, 412-415.

Shimura, Y., Saito, K., Imaizumi, S. and Yamamuro, C. (1996) Interrelationships Between Prosody and Gesture in Vocal Communication in the Early Years: Emotion Perception and Speaker Identification. Emergence of Human Cognition and Language. 3, 225-229.

Shimura, Y. and Imaizumi, S. (1996) Listener and context dependency in the perception of emotional aspects of infant voice. Acta Paediatrica Japonica. 38, 648-656.

Shriberg, E., Bates, R., Taylor, P., Stolcke, A., Jurafsky, D., Ries, K., Coccaro, N., Martin, R., Meteer, M. and Van Ess-Dykema, C. (1998) Can prosody aid the automatic classification of dialog acts in conversational speech? Language and Speech: Special Double Issue on Prosody and Conversation. 41(3-4), 443-492.

Smith, I.M. and Bryson, S.E. (1994) Imitation and action in autism: A critical review. Psychological Bulletin. 116 (2), 259-273.

Spangler, G., Geserick, B. and Von Wahlert, A. (2005) Parental perception and interpretation of infant emotions: Psychological and physiological processes. Infant and Child Development. 14 (4), 345-363.

Spence, M.J. and Freeman, M.S. (1996) Newborn infants prefer the maternal low-pass filtered voice, but not the maternal whispered voice. Infant Behaviour and Development. 19, 199-212.

Stern, D. N. (1974b) The goal and structure of mother-infant play. Journal of the American Academy of Child Psychiatry. 13, 402-421;

Stern, D. N. and Gibbon, J. (1980) Temporal expectancies of social behaviours in mother-infant play. In E. Thoman (Ed), Origins of Infant Social Responsiveness p. 409-429. New York, Erlbaum.

Stern, D.N., Spieker, S. and MacLean, K. (1982) Intonation contours as signals in maternal speech to prelinguistic infants. Developmental Psychology. 18 (5), 727-735.

Stern, D.N. (1985) The Interpersonal World of the Infant: A View from Psychoanalysis and Developmental Psychology. New York, Basic Books.

Stern, D.N., Hofer, L., Haft, W. and Dore, J. (1985) Affect attunement: The sharing of

feeling states between mother and infant by means of inter-modal fluency. In T.M. Field and N.A. Fox (Eds), Social Perception in Infants. p.249-268. Norwood, N.J. Ablex.

Stern, D. N. (1999). Vitality contours: The temporal contour of feelings as a basic unit for constructing the infant's social experience. In Rochat, P. (Ed), Early Social Cognition: Understanding Others in the First Months of Life p. 67-90. Mahwah, NJ, Erlbaum.

Stern, D. N. (2000) The Interpersonal World of the Infant: A View from Psychoanalysis and Development Psychology. (Second Edition) New York, Basic Books.

Stevenson, M.B., Ver Hove, J.N., Roach, M.A. and Leavitt, L.A. (1986) The beginning of conversation: Early patterns of mother-infant vocal responsiveness. Infant Behaviour and Development. 9 (4), 423-440.

Stormark, K.M. and Braarud, H.C. (2004) Infants' sensitivity to social contingency: A "double video" study of face-to-face communication between 2- and 4-month-olds and their mothers. Infant Behaviour and Development. 27, 195-203.

Studdert-Kennedy, M. (2002) Vocal Imitation, Facial Imitation, and the Gestural Origin of Linguistic Discrete Infinity. (Paper presentation). Evolution of Language, Fourth International Conference Plenary Talks. Harvard University. March.

Sullivan, M.W. and Lewis, M. (2003) Emotional expression of young infants and children: A practitioner's primer. Infants and Young Children. 16 (2), 120-142.

- Suzuki, M. J. (2000) Child-rearing and educational practices in the United States and Japan: Comparative perspectives. In: University of Hyogo Teacher Education Journal. 20 (1), 177-186.
- Takada, A. (2002) Mother-Infant Interactions Among the !Xun: Analysis of Gymnastic and Breast-Feeding Behaviours. (Paper presentation). Ninth International Conference on Hunting and Gathering Societies. Edinburgh, Heriot Watt University. September,
- Tincoff, R. and Jusczyk, P.W. (1999) Some beginnings of word comprehension in 6-month-olds. Psychological Science. 10 (2), 172-175.
- Tomasello, M. (1988) The role of joint attentional processes in early language development. Language Sciences. 10, 69-88.
- Tomasello, M., Kruger, A.C. and Ratner, H.H. (1993) Cultural learning. Behavioural and Brain Sciences. 16 (3), 495-511.
- Trainor, L.J., Austin, C.M. and Desjardins, R.N. (2000) Is infant-directed speech prosody a result of the vocal expression of emotion? Psychological Science. 11 (3), 188-195.
- Trehub, S.E.; Schellenberg, E.G.; Glenn, E. and Hill, D.S. (1997) The origins of music perception and cognition: A developmental perspective. In I. Deliège and J. Sloboda (Eds), Perception and Cognition of Music p.121-162. Hove, Erlbaum Psychology Press.
- Trehub, S.E., Unyk, A., Kamenetsky, S.B. and Hill, D.S. (1997) Mothers' and fathers'

- singing to infants. Developmental Psychology. 33(3), 500-507.
- Trehub, S. and Nakata, T. (2002) Emotion and music in infancy. Musicae Scientiae. Special Issue 2001-2002, 37-61.
- Trevarthen, C. and Hubley, P.A. (1978) Secondary intersubjectivity: Confidence, confiding and acts of meaning in the first year. In A. Lock (Ed), Action, Gesture and Symbol p. 183-229. London, Academic Press.
- Trevarthen, C. (1979) Communication and cooperation in early infancy: A description of primary intersubjectivity. In M.M Bullowa (Ed), Before Speech: The Beginning of Interpersonal Communication p.321-347. New York, Cambridge University Press.
- Trevarthen, C., Murray, L. and Hubley, P.A. (1981) Psychology of infants. In J. Davis and J. Dobbing (Eds), Scientific Foundations of Clinical Paediatrics (Second Edition) p. 211-274. London, Heinemann Medical.
- Trevarthen, C. (1993a) The function of emotions in early infant communication and development. In J. Nadel and L. Camaioni (Eds), New Perspectives in Early Communicative Development p.48-81. London, Routledge.
- Trevarthen, C. (1993b) The Self Born in Intersubjectivity: The Psychology of an Infant Communicating. In U. Neisser (Ed), The Perceived Self p.121-173. New York, Cambridge University Press.
- Trevarthen, C. (1994) Infant Semiosis. In W. Noth (Ed), Origins of Semiosis: Sign Evolution in Nature and Culture p.219-259. Berlin, Mouton de Gruyter.

- Trevarthen, C. and Aitken, K. (1994) Brain development, infant communication and empathy disorders: Intrinsic factors in child mental health. Development and Psychopathology. 6, 597-633.
- Trevarthen, C. (1998). The concept and foundations of infant intersubjectivity. In S. Bråten, (Ed), Intersubjective Communication and Emotion in Early Ontogeny p. 15-46. Cambridge, Cambridge University Press.
- Trevarthen, C. (1999). Musicality and the intrinsic motive pulse: Evidence from human psychobiology and infant communication. Musicae Scientiae, Special Issue, 1999-2000, 157-213.
- Trevarthen, C., Kokkinaki, T. and Fiamenghi, G. A. Jr. (1999). What infants' imitations communicate: With mothers, with fathers and with peers. In J. Nadel and G. Butterworth (Eds), Imitation in Infancy p. 127-185. Cambridge, Cambridge University Press.
- Trevarthen, C. (2001b) The neurobiology of early communication: Intersubjective regulations in human brain development. In A. F. Kalverboer & A. Gramsbergen (Eds), Handbook on Brain and Behavior in Human Development p.841-842. Dordrecht, Kluwer.
- Trevarthen, C. and Aitken K. J. (2001) Infant intersubjectivity: Research, theory and clinical applications. Annual Research Review, Journal of Child Psychology and Psychiatry. 42(1), 3-48.
- Trevarthen, C. and Reddy, V. (2007). Consciousness in infants. In M. Velman & S.

Schneider (Eds), A Companion to Consciousness p.41-57. Oxford, Blackwells.

Tronick, E.Z. (1989) Emotions and emotional communication in infants. American Psychologist. 44 (2), 112-119.

True, M.M., Pisanni, L. and Oumar, F. (2001) Infant-mother attachment among the Dogon of Mali. Child Development. 72 (5), 1451-1466.

Tsang, C. D. and Trainor, L. (2002) Spectral slope discrimination in infancy: sensitivity to socially important timbres. Infant Behaviour and Development. 25(2), 183 – 194.

Tsao, F., Liu, H. and Kuhl, P. (2006) Perception of native and non-native affricate-fricative contrasts: Cross-language tests on adults and infants. Journal of Acoustical Society of America. 120 (4), 2285-2294.

Tzourio-Mazoyer, N., De Schonen, S., Crivello, F., Reutter, B., Aujard, Y. and Mazoyer, B. (2002) Neural correlates of women face processing by 2-month-old infants. NeuroImage. 15, 454-461.

Tzourio-Mazoyer N, De Schonen S, Crivello F, Reutter B, et al. (2002) Neural correlates of woman face processing by 2-month-old infants. NeuroImage 15, 454–461.

Van Egeren, L.A., Barratt, M.S. and Roach, M.A (2001) Mother-infant responsiveness: Timing, mutual regulation, and interactional context. Developmental Psychology. 37 (5), 684-697.

- Valsiner, J. (1988) The Developing Body and Mind Number 6: Developmental Psychology in the Soviet Union. Sussex, The Harvester Press.
- Vandenberg, B. (1999) Infant communication. North American Journal of Psychology. 1(1), 29-34.
- Vygotsky, L. (1962) The Genetic Roots of Thought and Speech. Cambridge, M.A., MIT Press.
- Weisner, T.S. (2001) Childhood: Anthropological aspects. In N. J. Smelser and Paul B. Baltes (Eds), International Encyclopedia of the Social and Behavioural Sciences, Volume 3 p.1697-1701. Amsterdam, Elsevier.
- Werker, J.F. and Tees, R. (1984) Cross-language speech perception: Evidence for perceptual reorganisation during the first year of life. Infant Behaviour and Development. 7 (1), 49-63.
- Werker, J.F., Pons, F., Dietrich, C., Kajikawa, S., Fais, L. and Amano, S. (2007) Infant-directed speech supports phonetic category learning in English and Japanese. International Journal of Cognitive Science. 103, 1-147.
- Whittmann, M. and Poppel, E. (2000) Temporal mechanisms of the brain as fundamentals of communication with special reference to music perception and performance. Musicae Scientiae. Special Issue 1999-2000, 13-28.
- Wijnroks, L. (1997) Mother-infant interaction and contingency learning in pre-term infants. Early Development and Parenting. 6 (1), 27-36.
- Yale, M.E., Messinger, D.S., Cobo-Lewis, A.B., Oller, D.K. and Eilers, R.E.

- (1999) An event-based analysis of the coordination of early infant vocalisations and facial actions. Developmental Psychology. 35 (2), 505-513.
- Yamanouchi, I., Fukuhara, H. and Shimura, Y. (1990) The transmission of ambient noise and self-generated sound in the human body. Acta Paediatrica Japan. 32, 615-624.
- Zahavi, A. (2002) Messages and patterns in the vocal communication of babblers. In Evolution of Language. Fourth International Conference Plenary Talks. Paper presentation, Harvard University, Boston. March.
- Zeedyk, S. (2004) Imitation and Intimacy. Pre-Conference Workshop: Imitation and Communication. BPS Developmental Section Conference. Leeds, UK. September.
- Zeifman, D.M. and College, V. (2003) Predicting adult responses to infant distress: adult characteristics associated with perceptions, emotional reactions and timing of intervention. Infant Mental Health Journal. 24 (6), 597-612.